

Exhibit B

Report of the
APA Task Force on

Mental Health and Abortion

APA Task Force on Mental Health and Abortion

Brenda Major, PhD, Chair

Mark Appelbaum, PhD

Linda Beckman, PhD

Mary Ann Dutton, PhD

Nancy Felipe Russo, PhD

Carolyn West, PhD

A copy of the report is available online at <http://www.apa.org/pi/wpo/mental-health-abortion-report.pdf>

Suggested bibliographic reference: American Psychological Association, Task Force on Mental Health and Abortion. (2008). Report of the Task Force on Mental Health and Abortion. Washington, DC: Author. Retrieved from <http://www.apa.org/pi/wpo/mental-health-abortion-report.pdf>

Also APA reports synthesize current psychological knowledge in a given area and may offer recommendations for future action. They do not constitute APA policy nor commit APA to the activities described therein.

Copyright © 2008 by the American Psychological Association. This material may be reproduced in whole or in part without fees or permission provided that acknowledgment is given to the American Psychological Association. This material may not be reprinted, translated, or distributed electronically without prior permission in writing from the publisher. For permission, contact APA, Rights and Permissions, 750 First Street, NE, Washington, DC 20002-4242.

Printed in the USA.

REPORT OF THE APA TASK FORCE ON MENTAL HEALTH AND ABORTION

Table of Contents

Executive Summary	3
Introduction	5
Overview	6
Definitions and Scope of Report	6
Questions Addressed	7
Importance of Recognizing Variability in the Abortion Experience	9
Conceptual Frameworks	10
Abortion Within a Stress-and-Coping Perspective	10
Abortion as a Traumatic Experience	11
Abortion Within a Sociocultural Context	11
Abortion and Co-Occurring Risk Factors	12
Systemic risk factors	13
Personal risk factors	13
Summary of Conceptual Frameworks	14
Methodological Issues in Abortion Research	15
Comparison/Contrast Groups	15
Co-Occurring Risk Factors	16
Sampling	16
Measurement of Reproductive History and Problems of Underreporting	17
Attrition	18
Outcome Measures: Timing, Source, and Clinical Significance	18
Other Statistical Issues	19
Interpretational Problems and Logical Fallacies	19
Summary of Methodological Issues	20
Review of Scientific Literature	21
Search Strategy and Criteria for Inclusion	21
Descriptive Overview of Literature Identified for This Review	21
Review of Comparison Group Studies	22
Record-Based Studies and Secondary Analyses With Comparison Groups	22
Medical records	22
Secondary analyses of survey data	31
Evaluation of record-based and secondary analyses studies	55
Summary of medical-record and secondary analyses studies	56

Comparison Group Studies: Primary Data	57
Description of findings: U.S. samples	57
Description of findings: Non-U.S. samples	57
Evaluation of primary data comparison group studies	64
Studies of Abortion for Reasons of Fetal Abnormality	68
Description of findings	69
Evaluation of fetal abnormality studies	70
Review of Abortion-Only Studies	72
Prospective Studies	77
Retrospective Studies	87
Summary and Evaluation of Abortion-Only Studies	87
Summary and Conclusions	87
The Relative Risks of Abortion Compared to its Alternatives	87
Prevalence of Mental Health Problems Among Women in the United States Who Have Had an Abortion	90
Predictors of Individual Variation in Responses Following Abortion	91
Conclusions and Future Research	92
Endnotes	94
References	95
List of Tables	
Table 1A: Medical-Record Studies: U.S. Samples	104
Table 1B: Medical-Record Studies: International Samples	27
Table 2: Secondary Analyses of Survey Data: U.S. Samples and International Samples	39
Table 3A: Primary Data Comparison Group Studies: U.S. Samples	59
Table 3B: Primary Data Comparison Group Studies: International Samples	62
Table 4: Abortion for Reasons of Fetal Anomaly	70
Table 5: U.S. Samples of Abortion Group(s) Only: No Comparison Groups	75
Table 6: Population Estimates of Proportion of all Women and Women Identified as Having Been Pregnant Exceeding CES-D Clinical Cutoff Score, National Longitudinal Survey of Youth: 1992	91
Acknowledgments	105

REPORT OF THE APA TASK FORCE ON MENTAL HEALTH AND ABORTION

EXECUTIVE SUMMARY

The Council of Representatives of the American Psychological Association charged the Task Force on Mental Health and Abortion (TFMHA) with “collecting, examining, and summarizing the scientific research addressing the mental health factors associated with abortion, including the psychological responses following abortion, and producing a report based upon a review of the most current research.” In considering the psychological implications of abortion, the TFMHA recognized that abortion encompasses a diversity of experiences. Women obtain abortions for different reasons; at different times of gestation; via differing medical procedures; and within different personal, social, economic, and cultural contexts. All of these may lead to variability in women’s psychological reactions following abortion. Consequently, global statements about the psychological impact of abortion on women can be misleading.

The TFMHA evaluated all empirical studies published in English in peer-reviewed journals post-1989 that compared the mental health of women who had an induced abortion to the mental health of comparison groups of women ($N=50$) or that examined factors that predict mental health among women who have had an elective abortion in the United States ($N=23$). This literature was reviewed and evaluated with respect to its ability to address four primary questions: (1) Does abortion cause harm to women’s mental health? (2) How prevalent are mental health problems among women in the United States who have had an abortion? (3) What is the relative risk of mental health problems associated with abortion compared to its alternatives (other courses of action that might be taken by a pregnant woman in similar circumstances)? And, (4) What predicts individual variation in women’s psychological experiences following abortion?

A critical evaluation of the published literature revealed that the majority of studies suffered from methodological problems, often severe in nature. Given the state of the literature, a simple calculation of effect sizes or count of the number of studies that showed an effect in one direction versus another was considered inappropriate. The quality of the evidence that produced those effects must be considered to avoid misleading conclusions. Accordingly, the

TFMHA emphasized the studies it judged to be most methodologically rigorous to arrive at its conclusions.

The best scientific evidence published indicates that among adult women who have an *unplanned pregnancy* the relative risk of mental health problems is no greater if they have a single elective first-trimester abortion than if they deliver that pregnancy. The evidence regarding the relative mental health risks associated with multiple abortions is more equivocal. Positive associations observed between multiple abortions and poorer mental health may be linked to co-occurring risks that predispose a woman to both multiple unwanted pregnancies and mental health problems.

The few published studies that examined women's responses following an induced abortion due to fetal abnormality suggest that terminating a wanted pregnancy late in pregnancy due to fetal abnormality appears to be associated with negative psychological reactions equivalent to those experienced by women who miscarry a wanted pregnancy or who experience a stillbirth or death of a newborn, but less than those who deliver a child with life-threatening abnormalities.

The differing patterns of psychological experiences observed among women who terminate an unplanned pregnancy versus those who terminate a planned and wanted pregnancy highlight the importance of taking pregnancy intendedness and wantedness into account when seeking to understand psychological reactions to abortion.

None of the literature reviewed adequately addressed the prevalence of mental health problems among women in the United States who have had an abortion. In general, however, the prevalence of mental health problems observed among women in the United States who had a single, legal, first-trimester abortion for non-therapeutic reasons was consistent with normative rates of comparable mental health problems in the general population of women in the United States.

Nonetheless, it is clear that some women do experience sadness, grief, and feelings of loss following termination of a pregnancy, and some experience clinically significant disorders, including depression and anxiety. However, the TFMHA reviewed no evidence sufficient to support the claim that an observed association between abortion history and mental health was caused by the abortion per se, as opposed to other factors.

This review identified several factors that are predictive of more negative psychological responses following

first-trimester abortion among women in the United States. Those factors included perceptions of stigma, need for secrecy, and low or anticipated social support for the abortion decision; a prior history of mental health problems; personality factors such as low self-esteem and use of avoidance and denial coping strategies; and characteristics of the particular pregnancy, including the extent to which the woman wanted and felt committed to it. Across studies, prior mental health emerged as the strongest predictor of postabortion mental health. Many of these same factors also predict negative psychological reactions to other types of stressful life events, including childbirth, and, hence, are not uniquely predictive of psychological responses following abortion.

Well-designed, rigorously conducted scientific research would help disentangle confounding factors and establish relative risks of abortion compared to its alternatives, as well as factors associated with variation among women in their responses following abortion. Even so, there is unlikely to be a single definitive research study that will determine the mental health implications of abortion "once and for all" given the diversity and complexity of women and their circumstances.

REPORT OF THE APA TASK FORCE ON MENTAL HEALTH AND ABORTION

INTRODUCTION

Although the U.S. Supreme Court legalized abortion in the United States more than 35 years ago (*Roe v. Wade*, 1973), it continues to generate enormous emotional, moral, and legal controversy. Over the last two decades, one aspect of this controversy has focused on the effects of abortion on women's mental health (Bazelon, 2007; Cohen, 2006; Lee, 2003). Public debate on this issue can be traced to 1987, when then-President Ronald Reagan directed then-Surgeon General C. Everett Koop to prepare a Surgeon General's report on the public health effects (both psychological and physical) of abortion. After conducting a comprehensive review of the scientific literature, Dr. Koop declined to issue a report; instead, he sent a letter to President Reagan on January 9, 1989, in which he concluded that the available research was inadequate to support any scientific findings about the psychological consequences caused by abortion (Koop, 1989a). In subsequent testimony before Congress, Dr. Koop stated that his letter did not focus on the physical health risks of abortion because "obstetricians and gynecologists had long since concluded that the physical sequelae of abortion were no different than those found in women who carried to term or who had never been pregnant" (Koop, 1989, p. 195). Dr. Koop also testified that although psychological responses following abortion can be "overwhelming to a given individual," the psychological risks following abortion were "miniscule" from a public health perspective (Koop, 1989b, p. 241).

Dr. Koop's letter and an unofficial draft of his report read into the Congressional Record were cited by both abortion opponents and proponents to claim both the presence and absence of scientific evidence showing a detrimental effect of abortion on women's mental health (see Wilmoth, deAlterii, & Bussell, 1992). Recognizing the importance of this issue, the American Psychological Association (APA) convened a panel of scientific experts in February 1989 to conduct a review of the scientific literature on psychological responses to abortion. The panel focused on studies with the most rigorous research designs, reporting findings on the psychological status of women who had legal, elective, first-trimester abortions in the United States. Based on their review of this literature, the task force concluded that the most methodologically sound studies indicated that "severe negative re-

actions after legal, nonrestrictive, first-trimester abortion are rare and can best be understood in the framework of coping with a normal life stress" (Adler, David, Major, Roth, Russo, & Wyatt, 1990, pp. 43; see also Adler, David, Major, Roth, Russo, & Wyatt, 1992). The task force recognized that some individual women experience severe distress or psychopathology following abortion. However, the task force also noted that it was not clear that these symptoms are causally linked to the abortion. The conclusions of Dr. Koop and the 1989 APA Task Force have been widely regarded as the definitive scientific statements on the link between abortion and mental health.

Since publication of Koop's letter and unofficial draft report (1989a, 1989b) and the 1989 Task Force Report (see Adler et al., 1990), a number of new studies have been published in peer-reviewed journals that address the association between abortion and women's mental health. Some of these studies support the conclusions of the 1989 Task Force Report (e.g., Cohan, Dunkel-Schetter, & Lydon, 1993; Gilchrist, Hanaford, Frank, & Kay, 1995; Russo & Dabul, 1997; Russo & Zierk, 1992), whereas others challenge them (e.g., Cougle, Reardon, & Coleman, 2003; Fergusson, Horwood, & Ridder, 2006; Gissler, Kauppila, Merilainen, Toukomaa, & Hemminki, 1997; Reardon & Cougle, 2002a). Reviewers of this emerging literature have reached differing conclusions. Based on their review of the post-1990 literature, for example, Bradshaw and Slade (2003) concluded that "The conclusions drawn from the recent longitudinal studies looking at long-term outcomes following abortion, as compared to childbirth, mirror those of earlier reviews (e.g., Adler et al., 1992; Wilmoth et al., 1992), with women who have abortions doing no worse psychologically than women who give birth to wanted or unwanted children" (p. 948). In contrast, in testimony introduced in support of a law that would have banned all abortions in South Dakota except for those in which the mother's life was in danger, Coleman (2006b) concluded that the scientific evidence shows that "abortion poses significant risk to women's mental health and carries a greater risk of emotional harm than childbirth."

Recognizing the need for a critical review of the recent literature, in 2006 the Council of Representatives of APA established a new Task Force on Mental Health and Abortion (TFMHA) composed of scientific experts in the areas of stigma, stress and coping, interpersonal violence, methodology, women's health, and reproductive health. The APA Council charged the TFMHA with "collecting, examining, and summarizing the

scientific research addressing the mental health factors associated with abortion, including the psychological responses following abortion, and producing a report based upon a review of the most current research.”

This report summarizes the findings of the 2006 TFMHA. This report updates rather than duplicates efforts of the 1989 Task Force. We refer the reader to Adler et al. (1992) for a discussion of APA's involvement in abortion-related issues, the history and status of abortion in the United States, and a methodological critique of the literature on abortion prior to 1990 (see also the fall 1992 issue of the *Journal of Social Issues*).

In preparing this report, the TFMHA recognized that differing moral, ethical, and religious perspectives affect how abortion is perceived. Furthermore, it recognized that members of APA are likely to have a wide range of personal views on abortion. Irrespective of their views on the morality of abortion, however, APA members are united in valuing carefully and rigorously collected and interpreted scientific evidence. The TFMHA considered its mission not only to review, but also to critically and objectively evaluate the quality of the scientific evidence without regard to the direction of its findings in order to draw conclusions about the mental health implications of abortion based on the best scientific evidence available. This TFMHA report represents the most thorough, current, and critical evaluation of the literature published since 1989 (see Bradshaw & Slade, 2003; Coleman, Reardon, Strahan, & Cogle, 2005; Dagg, 1991; Posavac & Miller, 1990; Stotland, 1997; Thorp, Hartmann, & Shadigian, 2003, for prior published reviews of this literature).

Overview

We begin this report by defining terms, outlining the scope of the TFMHA report, and specifying the questions that the research literature has been used to address (Section I). Next, we discuss conceptual frameworks important for understanding the empirical literature on abortion and mental health (Section II) and important methodological issues to consider in evaluating this literature (Section III). We then turn to the core of our report (Sections IV and V): a review and evaluation of empirical studies published in English in peer-reviewed journals post-1989 that compares the mental health of women who have had an elective abortion to the mental health of various comparison groups (see detailed inclusion criteria below). We reviewed only peer-reviewed studies in order to include only research findings that stood the test of in-

dependent scrutiny of qualified scientific experts. In a following section (Section VI), we review research published post-1989 in the United States that has addressed factors that predict mental health among women who have had an elective abortion. We end with a summary and conclusions based on our review (Section VII).

Definitions and Scope of Report

There are multiple ways to conceptualize the mental health implications of abortion and many empirical literatures that are relevant to this topic. Studies examining the mental health implications of childbearing, particularly of unwanted childbearing, or of single parenting, for example, are relevant for comparison purposes (see Barber, Axinn, & Thornton (1999) for information on mothers with unwanted births). So, too, are studies of the effects on children of being born unwanted (see David, Dytrych, & Matejcek, 2003) or on women of being denied abortion (see Dagg, 1991). To review all of those literatures in this report, however, would be a massive undertaking beyond the scope and charge of this task force. To keep its task manageable, the TFMHA limited its review and evaluation to the empirical literature on the implications of induced or intentional termination of pregnancy for women's mental health. We do not consider the implications of abortion for the mental health of fathers, other children or family members, or clinic workers. Although these are important questions worthy of study, they are beyond the scope of this report.

Our review is limited to studies examining the mental health implications of *induced abortion*. In some studies, induced termination of pregnancy is not differentiated from spontaneous termination of pregnancy (spontaneous abortion, or miscarriage). Although spontaneous abortion may have mental health consequences, we consider those consequences only when they are compared with those of induced abortion. Other terms used to indicate induced abortion include *elective abortion*, *voluntary abortion*, and *therapeutic abortion*. These distinctions can be important. Given that abortion involves a medical procedure, the term therapeutic would seem to apply to all abortions. However, typically the term is applied to abortions induced for medically related reasons, such as to protect the mother's health or because of severe fetal abnormalities. This term also was used to describe abortions

performed for psychiatric reasons prior to legalization of abortion in the United States. Almost all abortions (92% according to the 2002 National Survey of Family Growth) in the United States are of unintended pregnancies, pregnancies that are not induced for therapeutic reasons (Finer & Henshaw, 2006a). A late-term induced abortion of an intended pregnancy may have very different implications for mental health than a first-trimester induced abortion of an unintended pregnancy.

We also limited our review to studies examining the implications of induced abortion for *mental health outcomes*. Other outcomes potentially related to abortion (either as antecedents or consequences), such as education, income, occupational status, marital status, and physical health, are beyond the scope of this report. We conceptualized mental health broadly, relying on the World Health Organization (WHO) definition of mental health as a “state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community” (World Health Organization [WHO], 2007). This report thus considers a wide array of outcomes related to mental health, including measures of psychological well-being (e.g., self-esteem, life satisfaction), emotions (e.g., relief, sadness), problem behaviors (e.g., substance abuse, child abuse), and measures of severe psychopathology. In considering the mental health implications of abortion, it is crucial to distinguish between clinically significant mental disorders, such as major depression, generalized anxiety disorder, or posttraumatic stress disorder, and a normal range of negative emotions or feelings one might experience following a difficult decision, such as feelings of regret, sadness, or dysphoria. While the latter feelings may be significant, by themselves they do not constitute psychopathology. In this report, we use the term *mental health problems* to refer to clinically significant disorders assessed with valid and reliable measures or physician diagnosis. We use the term *negative psychological experiences or reactions* to refer to negative behaviors (e.g., substance use) and emotions (e.g., guilt, regret, sadness), and the term *psychological well-being* to refer to positive outcomes, such as self-esteem and life satisfaction. Because most studies published during the review period framed their research in terms of mental health problems and the

negative experiences or reactions of women, this report, of necessity, emphasized these outcomes rather than psychological well-being following abortion.

Our core review and evaluation was also limited to studies that met the following inclusion criteria: (a) empirical research, (b) published in English, (c) in peer-reviewed journals, (d) subsequent to 1989, (e) measuring a mental health relevant outcome subsequent to abortion, and (f) including a comparison group of women (see details on selection criteria, below).

In addition to these core studies, the TFMHA reviewed studies based on U.S. samples that met the above inclusion criteria but did not include a comparison group of women. Because such studies do not include a comparison group, they cannot be used to draw conclusions about relative risks of abortion compared to its alternatives. Nonetheless, these studies provide important insight into sources of variability in women’s experiences of abortion in the U.S. context.

Questions Addressed

When considering the empirical literature on the association between abortion and mental health, it is useful to keep in mind four primary questions that this literature addressed: (1) Does abortion cause harm to women’s mental health? (2) How prevalent are mental health problems among women in the United States who have had an abortion? (3) What is the relative risk of mental health problems associated with abortion compared to its alternatives (other courses of action that might be taken by a pregnant woman in similar circumstances)? And, (4) What predicts individual variation in women’s psychological experiences following abortion? As we discuss below, each of these different questions requires a different research approach. Some of these questions are scientifically testable; others are not.

Does abortion cause harm to women’s mental health?

Although this is the question that is posed most often in public debates, this question is not scientifically testable as stated. An adequate answer to this question requires a randomized experimental design that would rigorously define the experimental, control, and outcome variables and specify any limitations in generalizing the results. Unlike many other

areas of research, however, the study of abortion is not open to the methodologies of randomized clinical trials. For obvious reasons, it is neither desirable nor ethical to randomly assign women who have unwanted pregnancies to an abortion versus delivery versus adoption group. Thus, although people have frequently used the existing literature to make causal statements, inferences of cause from this literature are inappropriate.

How prevalent are mental health problems among women in the United States who have had an abortion? This question focuses attention on the extent to which abortion poses a threat to women's mental health, i.e., is associated with a clinically significant mental disorder (see Wilmoth et al., 1992 for a discussion of this issue). In order to answer this question, research must have several key characteristics. First, the research must be based on samples of women representative of the women to whom one wants to generalize. Thus, to address whether abortion poses a threat to the mental health of women in the United States requires a study based on a nationally representative sample of women in the United States. Highly selected samples, biased samples, samples with considerable attrition or underreporting, or samples of women in other cultures and social contexts are not appropriate for answering this question. As will be discussed below, sampling problems are a serious concern in abortion research. Second, an adequate answer to the prevalence question also requires a clearly defined and agreed-upon definition of a "mental health problem" and a valid, reliable, and agreed-upon measurement of that problem. Feelings of sadness or regret within the normal range of emotion are not clearly defined and agreed-upon mental health problems. Mental health outcomes that meet established criteria for clinically significant disorders are. Third, researchers must know the prevalence of the same mental health problem in the general population of U.S. women who share characteristics similar to the abortion group, e.g., women who are of a similar age and demographic profile. Such information is essential for interpreting the significance of findings. For example, if 15% of women in a nationally representative sample who had had an abortion were found to meet diagnostic criteria for depression, the meaning of this would be more a cause for concern if the base rate for clinical depression among women in the general population of a similar age and demographic profile was 5% than if it was 25%.

What is the relative risk of mental health problems associated with abortion compared to its alternatives (other courses of action that might be taken by a pregnant woman in similar circumstances)?

This question addresses *relative risk*. It focuses attention on the crucially important but frequently overlooked point that the outcomes associated with elective abortion must be compared with the outcomes associated with other courses of action that might be taken by a pregnant woman in similar circumstances (i.e., facing an unwanted pregnancy). *Once a woman is pregnant, there is no mythical state of "unpregnancy."* Questions of relative risk include: How does the mental health of a woman who has an abortion compare to the mental health that a woman in comparable circumstances would experience were she not to have an abortion or were she to be denied an abortion? Are negative feelings that may accompany abortion of an unwanted pregnancy more severe than alternative solutions, such as giving up a child for adoption or raising a child a woman does not want or feels emotionally, physically, or financially unable to care for? Only research designs that include a comparison group that is clearly defined and otherwise equivalent to women who have an elective abortion are appropriate for answering this primary question. Otherwise, any previously existing group differences associated with the outcome variable may bias conclusions. As will be discussed below, few studies examining the mental health implications of abortion include appropriate comparison groups for answering this question.

What predicts individual variation in women's psychological experiences following abortion? This last question addresses the substantial individual variation observed in women's psychological experiences following abortion. Rather than focusing on how the "typical" woman responds following a "typical" abortion, this question asks why some women experience abortion more or less favorably than do others. This question is important to address because the proportion of women who have negative mental health issues after having an abortion will vary depending on the characteristics of each woman as well as the characteristics of her circumstances—there is no one answer that applies to all women. Because this question focuses on within-group variability rather than on differences between the abortion group and another group, research designed to answer this question does not require a comparison group of women who do not have

abortions, or a nationally representative sample. Research designed to answer this question, however, should at minimum be prospective and longitudinal and use reliable and valid measures of mental health.

Variability in the Abortion Experience

In considering the psychological implications of abortion, it is important to recognize that the term abortion encompasses a diversity of experiences and means different things to different women. Women obtain abortions for a variety of reasons, at different times of gestation, via differing medical procedures, all of which may affect the experience of abortion. Women's responses after abortion do not only reflect the meaning of abortion to her; they also reflect the meaning of pregnancy and motherhood, which varies among women. Furthermore, women obtain abortions within widely different personal, social, economic, religious, and cultural contexts that shape the cultural meanings and associated stigma of abortion and motherhood as well as others' responses to women who have abortion. All of these may lead to variability in women's psychological experiences to their particular abortion experience. For these reasons, global statements about the psychological impact of abortion on women can be misleading.

Women obtain abortions for different reasons. The vast majority of abortions are of unintended pregnancies—either mistimed pregnancies that would have been wanted at an earlier or later date or unwanted pregnancies that were not wanted at that time or at any time in the future (Henshaw, 1998; Torres & Forrest, 1988). Approximately half of women in the United States will face an unintended pregnancy during their lifetime, and about half of those who unintentionally become pregnant resolve the pregnancy through abortion (Finer & Henshaw, 2006a). The reasons that women most frequently cite for terminating a pregnancy include not being ready to care for a child (or another child) at that time, financial inability to care for a child, concern for or responsibility to others (especially concerns related to caring for a future child and/or for existing children), desire to avoid single parenthood, relationship problems, and feeling too young or immature to raise a child (Finer, Frowirth, Dauphinee, Singh, & Moore, 2005). Some pregnancies are terminated because they are a consequence of rape or incest. Very few (<1%) women cite coercion from others as a major reason for their abortion (Finer

et al., 2005). A very small percentage of abortions are of planned and wanted pregnancies. Women who terminate wanted pregnancies typically do so because of fetal anomalies or risks to their own health.

Gestational age at time of abortion varies. The vast majority (over 90%) of abortions in the United States occur in the first trimester of pregnancy (Boonstra, Gold, Richards, & Finer, 2006). Later-trimester abortions occur for a variety of reasons. In some cases, particularly involving teenagers, a woman may be unaware that she is pregnant until the second trimester or must go through legal proceedings (e.g., judicial bypass) in order to obtain an abortion (Boonstra et al., 2006). Later-trimester abortions also are performed after discovery of fetal abnormalities or risks to the mother's health.

Abortion procedures vary as well. Although most first-trimester abortions are performed using electric vacuum aspiration (EVA), nonsurgical methods involving use of a drug or combination of drugs to terminate pregnancy (e.g., mifepristone) are increasingly being used. Nonsurgical abortions comprised 14% of nonhospital abortions in 2005 as compared to 6% in 2001 (Jones, Zolna, Henshaw, & Finer, 2008). Procedures for abortions later than the first trimester include dilation and evacuation and induction of labor.

The experience of abortion may also vary as a function of a woman's ethnicity and culture. The United States is home to a growing number of ethnic and immigrant populations, including Hispanic (13%), African American (12.9%), and Asian and Pacific Islanders (4.2%). According to the 2000 Census data, African American women are more than three times as likely as White women to have an abortion (Dugger, 1998). Latinas are approximately two times as likely as White women to have an abortion, although there are important subgroup differences. Based on estimates from the Hispanic Health and Nutrition Examination Survey, among Latinas, Mexican women used abortion least; Puerto Rican women used abortion more than Mexican women, and Cubans used abortion the most (Erickson & Kaplan, 1998). The overrepresentation of ethnic minority women among those who obtain abortions in the United States may represent the general problem of greater poverty and reduced access to health care, including reproductive health services, among women of color. Although

there appears to be a strong influence of traditional African American and Latino cultural and religious values on women's use of abortion, this influence varies by age, country or area of ancestry or origin, level of acculturation, socioeconomic status, and educational and occupational attainment (Dugger, 1998; Erickson & Kaplan, 1998). Thus, it appears that for women of color, moral and religious values intersect with identities conferred by race, class, or ethnicity to influence women's likelihood of obtaining an abortion and, potentially, their psychological experiences following it. Historical linkages between coercive abortion and sterilization practices and the eugenics movements may lead some poor women and women of color to feel ambivalent on the issue of abortion despite understanding the importance of reproductive choice (Dugger, 1998; Erickson & Kaplan, 1998).

Women's experience of abortion may also vary as a function of the developmental phase of the life cycle in which it occurs. A teenager who terminates her first pregnancy, for example, may experience different psychological effects compared to an adult woman who terminates a pregnancy after giving birth to several children.

Women's experience of abortion may also vary as a function of their religious, spiritual, and moral beliefs and those of others in their immediate social context. There are religious denominational differences in social attitudes toward abortion (e.g., Bolzendahl & Brooks, 2005). Women who belong to religious groups that oppose abortion on moral grounds, such as Evangelical Protestants or Catholics, may be more conflicted about terminating a pregnancy through abortion. Religiosity and religious beliefs are likely to shape women's likelihood of having an abortion, as well as their responses to abortion.

In summary, women's psychological experience of abortion is not uniform, but rather varies as a function of characteristics and events that led up to the pregnancy; the circumstances of women's lives and relationships at the time that a decision to terminate the pregnancy was made; the reasons for, type, and timing of the abortion; events and conditions that occur in women's lives subsequent to an abortion; and the larger social-political context in which abortion takes place. This variability is an important factor in under-

standing women's psychological experiences following abortion.

CONCEPTUAL FRAMEWORKS

Much of the research examining the psychological implications of abortion has been atheoretical (Posavac & Miller, 1990). Nonetheless, several different perspectives have shaped understanding of potential associations between abortion and mental health outcomes. These perspectives are not necessarily mutually exclusive and are often complementary. Yet, they lead to different questions and different methodological approaches and can lead to different conclusions.

Abortion Within a Stress and Coping Perspective

One frequently used framework for understanding women's psychological experience of abortion is derived from psychological theories of stress and coping (e.g., Lazarus & Folkman, 1984). This perspective views abortion as a potentially stressful life event within the range of other normal life stressors (Adler et al., 1990, 1992). Because abortion occurs in the context of a second stressful life event—a pregnancy that is unwanted, unintended, or associated with problems in some way—a stress and coping perspective emphasizes that it can be difficult to separate out psychological experiences associated with abortion from psychological experiences associated with other aspects of the unintended pregnancy (Adler et al., 1990, 1992). Abortion can be a way of resolving stress associated with an unwanted pregnancy, and, hence, can lead to relief. However, abortion can also engender additional stress of its own.

A hallmark principle of psychological theories of stress and coping is variability (e.g., Billings & Moos, 1981; Lazarus & Folkman, 1984). From this perspective, although unwanted pregnancy and abortion can pose challenges and difficulties for an individual woman, these events will not inevitably or necessarily lead to negative psychological experiences for women. Stress emerges from an interaction between the person and the environment in situations that the person appraises as taxing or exceeding his or her resources to cope. A woman's psychological experience of abortion will be mediated by her appraisals of the pregnancy and abor-

tion and their significance for her life, her perceived ability to cope with those events, and the ways in which she copes with emotions subsequent to the abortion. These are shaped by conditions of the woman's environment (e.g., age, material resources, presence or absence of a supportive partner) as well as by characteristics of the woman herself (e.g., her personality, attitudes, and values). Thus, for example, a woman who regards abortion as conflicting with her own and her family's deeply held religious, spiritual, or cultural beliefs but who nonetheless decides to terminate an unplanned or unwanted pregnancy may appraise that experience as more stressful than would a woman who does not regard an abortion as in conflict with her own values or those of others in her social network.

Research derived from a stress-and-coping perspective has identified several factors that are associated with more negative psychological reactions among women who have had an abortion. These include terminating a pregnancy that is wanted or meaningful; perceived pressure from others to terminate a pregnancy; perceived opposition to the abortion from partners, family, and/or friends; and a lack of perceived social support from others. Other factors found to be associated with more negative postabortion experiences include personality traits (e.g., low self-esteem, a pessimistic outlook, low-perceived control) and a history of mental health problems prior to the pregnancy (see Adler et al., 1992; Major & Cozzarelli, 1992; Major et al., 2000 for reviews).

Importantly, many of the same individual and interpersonal factors that predict how women will appraise, cope with, and react psychologically to abortion are also predictors of how women will appraise, cope with, and react psychologically to other types of stressful life events, including unwanted motherhood or relinquishment of a child for adoption. For instance, low-perceived social support, low self-esteem, and pessimism also are risk factors for postpartum depression (Beck, 2001; Grote & Bledsoe, 2007; Logsdon & Usui, 2001). Consequently, the same risk factors for adverse reactions to abortion can also be risk factors for adverse reactions to its alternatives.

Abortion as a Traumatic Experience

Whereas the above framework views abortion within the range of normal life stressors, an alternative perspective views abortion as a uniquely traumatic experi-

ence. This perspective argues that abortion is traumatic because it involves a human death experience, specifically, the intentional destruction of one's unborn child and the witnessing of a violent death, as well as a violation of parental instinct and responsibility, the severing of maternal attachments to the unborn child, and unacknowledged grief (e.g., Coleman, Reardon, Strahan, & Cougle, 2005; MacNair, 2005; Speckhard & Rue, 1992). The view of abortion as inherently traumatic is illustrated by the statement that "once a young woman is pregnant...it is a choice between having a baby or having a traumatic experience" (original italics; Reardon, 2007, p. 3). The belief that women who terminate a pregnancy typically will feel grief, guilt, remorse, loss, and depression also is evident in early studies of the psychological implications of abortion, many of which were influenced by psychoanalytic theory and based on clinical case studies of patients presenting to psychiatrists for psychological problems after an abortion (see Adler et al., 1990).

Speckhard and Rue (1992; Rue, 1991, 1995) posited that the traumatic experience of abortion can lead to serious mental health problems for which they coined the term *postabortion syndrome* (PAS). They conceptualized PAS as a specific form of posttraumatic stress disorder (PTSD) comparable to the symptoms experienced by Vietnam veterans, including symptoms of trauma, such as flashbacks and denial, and symptoms such as depression, grief, anger, shame, survivor guilt, and substance abuse. Speckhard (1985, 1987) developed the rationale for PAS in her doctoral dissertation in which she interviewed 30 women specifically recruited because they deemed a prior abortion experience (occurring from 1 to 25 years previously) to have been "highly stressful." Forty-six percent of the women in her sample had second-trimester abortions, and 4% had third-trimester abortions; some had abortions when it was illegal. As noted above, this self-selected sample is not typical of U.S. women who obtain abortions. PAS is not recognized as a diagnosis in the *Diagnostic and Statistical Manual of the American Psychiatric Association* (American Psychiatric Association, 2002).

Abortion Within a Sociocultural Context

A third perspective emphasizes the impact of the larger social context within which pregnancy and abortion occur on women's psychological experience of these events. Unwanted pregnancy and abortion do

not occur in a social vacuum. The current sociopolitical climate of the United States stigmatizes some women who have pregnancies (e.g., teen mothers) as well as women who have abortions (Major & Gramzow, 1999). It also stigmatizes the nurses and physicians who provide abortions. From a sociocultural perspective, social practices and messages that stigmatize women who have abortions may directly contribute to negative psychological experiences post abortion.

The psychological implications of stigma are profound (see Major & O'Brien, 2005, for a review). Experimental studies have established that stigmatization can create negative cognitions, emotions, and behavioral reactions that can adversely affect social, psychological, and biological functioning. Effects of perceived stigma include cognitive and performance deficits (Steele & Aronson, 1995), increased alcohol consumption (Taylor & Jackson, 1990), social withdrawal and avoidance (Link, Struening, Rahav, Phelan, & Nuttbrock, 1997), increased depression and anxiety (Taylor, Henderson, & Jackson, 1991), and increased physiological stress responses (Blascovich, Spencer, Quinn, & Steele, 2001). Societal stigma is particularly pernicious when it leads to "internalized stigma"—the acceptance by some members of a marginalized group of the negative societal beliefs and stereotypes about themselves. Women who come to internalize stigma associated with abortion (e.g., who see themselves as tainted, flawed, or morally deficient) are likely to be particularly vulnerable to later psychological distress.

A sociocultural context that encourages women to believe that they "should" or "will" feel a particular way after an abortion can create a self-fulfilling prophecy whereby societally induced expectancies can become confirmed. Mueller and Major (1989) demonstrated experimentally the effect of expectancies on women's psychological experiences after abortion. They randomly assigned women prior to their abortion to one of three short counseling interventions. One intervention focused on improving women's self-efficacy for coping with abortion (creating positive coping expectancies), another focused on reducing the extent to which women attributed their pregnancy to their character (as opposed to their behavior), and the third focused on birth control. Women exposed to the self-efficacy intervention were significantly less likely

to display depressed affect following the abortion than those in the other two conditions. Societal messages that convey the expectation that women will cope poorly with an abortion would be expected to have the reverse effect; i.e., by creating negative coping expectancies, they may cause women to feel bad following an abortion.

Whether or not a particular behavior or attribute is stigmatized often varies across cultures and time (Crocker, Major, & Steele, 1998). Actions that once were viewed benignly can become stigmatized (e.g., smoking), and others that once were highly stigmatized (e.g., sex out of wedlock, divorce, cohabitation) can become less so. As society's views of a behavior change, so too will the appraisals and responses of those who engage in that behavior. Hence, the sociocultural context can shape a woman's appraisal of abortion not only at the time that she undergoes the procedure, but also long after the abortion. Social messages that encourage women to think about (reappraise) a prior abortion in more negative ways (as a sin, as killing a child) may increase women's feelings of guilt, internalized stigma, and emotional distress about an abortion they had long ago. In contrast, social messages and support groups that encourage women to cognitively reappraise an abortion in a more positive or benign way may lead to improved emotional responses (Trybulski, 2006).

Abortion and Co-Occurring Risk Factors

A fourth conceptual framework for understanding women's postabortion mental health emphasizes systemic, social, and personal factors that are precursors to unintended pregnancy and, hence, place women at risk for having abortions and/or predispose them to experience mental health problems, regardless of pregnancy and its resolution. From this perspective, mental health problems that develop after an abortion may not be caused by the procedure itself, but instead reflect other factors associated with having an unwanted pregnancy or antecedent factors unrelated either to pregnancy or abortion, such as poverty, a history of emotional problems, or intimate-partner violence. This co-occurring risk perspective emphasizes that aspects of a woman's life circumstances and psychological characteristics prior to or co-occurring with her pregnancy must be considered in order to make sense of any mental health problems observed subsequent to abortion.

Unwanted pregnancies are not random events. The lives of women who have unwanted pregnancies or abortions differ in a variety of ways from the lives of women who do not have unwanted pregnancies or abortions, and do so before, during, and after pregnancy occurs. These differences may have implications for later functioning apart from any influence from the experience of unwanted pregnancy and/or abortion. The necessity of considering preexisting or co-occurring group differences is widely recognized by researchers who study the consequences of nonmarital and adolescent births (e.g., Moore, 1995). As described below, substantial research literature has shown that systemic and personal characteristics that predispose women to have unintended pregnancies also predispose them to have psychological and behavioral problems. Consequently, correlations between abortion status and mental health problems observed after an abortion may be spurious due to their joint association with similar risk factors present prior to the pregnancy. We briefly review evidence consistent with this perspective below.

Systemic risk factors. Poverty is a systemic risk factor for unplanned pregnancy and for abortion. Women at particularly high risk for unintentional pregnancy and women who obtain abortions tend to be young, unmarried, poor, and women of color (Finer & Henshaw, 2006a; Jones, Darroch, & Henshaw, 2002a, 2002b; Jones & Kost, 2007). In 2000, women with resources below the federal poverty level represented 57% of all abortions (Jones, Darroch, & Henshaw, 2002b). Exposure to sexual or physical abuse during childhood and exposure to intimate partner violence including rape also are associated with greater likelihood for both unintended pregnancy and abortion (e.g., Boyer & Fine, 1992; Dietz et al., 1999; Gazmararian, Lazorick, Spitz, Ballard, Saltzman, & Marks, 1996; see Coker, 2007; Pallitto & O'Campo, 2005; Russo & Denious, 1998b for reviews).

From a co-occurring risks perspective, the greater exposure to adverse life circumstances (poverty, abuse, and intimate violence) among the group of women who have abortions compared with other women may underlie a positive correlation observed between abortion and mental health problems. Given the former's greater exposure to adversity, the absence of such an association would be noteworthy.

Indeed, these same systemic factors shown to be associated with increased risk for unintended pregnancy and abortion have also been shown to be associated with increased risk for mental health problems. For example, studies based on nationally representative samples show that poverty is strongly related to an increased likelihood of psychiatric disorder (e.g., Kessler et al., 1994; Robins & Regier, 1991). Children who grow up in poor neighborhoods are at higher risk for teen pregnancy, substance abuse, obesity, smoking, and dropping out of school, all of which are risk factors for psychological problems (Mather & Rivers, 2006; Messer, Kaufman, Dole, Savitz, & Laraia, 2006). Exposure to domestic (intimate) violence also is a strong and well-documented predictor of physical and mental health problems, including suicide, post-traumatic stress disorder, depression, and substance abuse (see Golding, 1999, for a meta-analysis and review). The more violence-related events a woman has experienced and the more stressful life events she has experienced in general, the greater her risk for developing a mental disorder (Breslau, Kessler, Chilcoat, Schultz, Davis, & Andreski, 1998; Brown & Harris, 1978; Golding, 1999).

Personal risk factors. In addition to systemic factors, personality or behavioral factors may also predispose a woman to unplanned pregnancy and abortion, as well as to mental health problems. There is substantial evidence that problem behaviors tend to co-occur among the same individuals. For example, high school students who report engaging in early sexual activity also are more likely to report smoking; using alcohol, marijuana, and hard drugs; minor delinquency; and, to a lesser extent, major aggression and gambling (Willoughby, Chalmers, & Busseri, 2004). Women who have unintended pregnancies and abortions are more likely than other women to have previously engaged in behaviors such as smoking, using alcohol and illicit drugs, engaging early in sexual intercourse, and having unprotected sexual intercourse (Costa, Jessor, & Donovan, 1987).

One explanation for this pattern is that involvement in problem behaviors follows definite pathways in which specific factors place the individual who has participated in one behavior (e.g., drug use) at risk of initiating another (e.g., early sexual activity), which puts that person at risk for another event (unintended pregnancy), which in turn puts that person at risk for

another event (abortion) (e.g., Kandel, 1989). A longitudinal study based on data from the National Longitudinal Study of Youth (NLSY) showed that drug use among young women greatly increased their risk of early sexual activity (before age 16) when other important risk factors were controlled (Rosenbaum & Kandel, 1990). In a subsequent study also based on data from the NLSY, Mensch and Kandel (1992) showed that drug use was uniquely predictive of both subsequent premarital teen pregnancy and the decision to terminate a premarital teen pregnancy. To avoid confounding antecedents of pregnancy with its consequences, they restricted their analyses to the youngest birth cohorts in the sample. This ensured that the measurement of the independent variables (e.g., drug use) preceded the events of interest (premarital teen pregnancy and abortion). They found that the risk of premarital teen pregnancy was nearly four times as high for women who had used illicit drugs other than marijuana as it was for women with no history of prior substance involvement. Furthermore, early illicit drug use was the strongest predictor of a later abortion. Another prospective longitudinal study found that women who at age 18 (none of whom had had a pregnancy or abortion) had reported smoking or using drugs had an increased likelihood of a subsequent unplanned pregnancy and, as a result, higher rates of abortion by age 29 compared to women who at age 18 had not reported using these drugs (Martino, Collins, Ellickson, & Klein, 2006).

An alternative explanation for the co-occurrence of problem behaviors is that individuals who engage in problem behaviors such as alcohol or drug use share a set of personality characteristics that predisposes them to engage in risky behaviors that increase the likelihood of other problems (e.g., unplanned pregnancy; Jessor & Jessor, 1977; see Dryfoos, 1990, for a review). For example, scoring high on a measure of "unconventionality" has been found to positively predict both abortion and unplanned pregnancy (Martino, Collins, Ellickson, & Klein, 2006). Personality factors that diminish a person's ability to regulate negative emotion may also put him or her at risk for engaging in problem behaviors. In a longitudinal study of a representative sample of 1,978 Black and White adolescents, Cooper, Wood, Orcutt, and Albino (2003) found that high impulsivity and an avoidance style of coping with negative emotions were risk factors for involvement in a wide range of problem behaviors, in-

cluding risky sexual behavior, substance use, delinquent behavior, and educational underachievement. Furthermore, an avoidance coping style prospectively predicted initial or increasing involvement in all of these problem behaviors among individuals with no prior experience with that behavior. Thus, for example, girls high in avoidance coping who had little or no prior sexual experience were subsequently more likely to engage in risky sexual behavior than girls lower in avoidance coping. Because early sexual activity and risky sexual behavior are risk factors for unintended pregnancy, which in turn is a risk factor for abortion, being high in avoidance styles of coping with negative emotion may be a predisposing risk factor for the experience of abortion.

Importantly, many of these personal characteristics that put women at risk for problem behaviors and unplanned pregnancy also put them at risk for mental or physical health problems, *whether or not a pregnancy is aborted or carried to term*. For example, a number of studies demonstrate that using avoidant forms of coping with negative emotions is associated with poorer mental health and exacerbates adjustment difficulties over time, even after controlling for prior levels of adjustment (Aldwin & Revenson, 1987; Major, Richards, Cooper, Cozzarelli, & Zubek, 1998). The best predictor of mental health problems later in life is a prior occurrence of mental health problems. For example, Kessler, Avenevoli, and Merikangas (2001) reported that 50% of adolescents who had an occurrence of major depression and 90% of adolescents who experienced mania during their adolescence continued to have recurrences of these disorders in adulthood.

Summary of Conceptual Frameworks

The four perspectives summarized above can be complementary ways of understanding underlying causes of women's psychological experience of abortion. The first perspective regards abortion as a stressful life event similar to other types of stressful life events a woman may experience. According to this perspective, women will vary markedly in how they appraise, cope with, and adjust to unwanted pregnancy and abortion, just as people vary widely in how they respond to other types of stressful life events. A stress-and-coping perspective thus does not rule out the possibility that some women may experience severe negative psychological experiences following abortion, but locates

such reactions in women's appraisals and coping processes and the personal and social factors that shape those, rather than in the nature of the event itself. In contrast, the second perspective suggests that due to its unique features, abortion is likely to be experienced as traumatic by most women. Thus, in contrast to other perspectives discussed, this particular framework suggests that most women will have negative psychological experiences subsequent to abortion.

The sociocultural perspective emphasizes that women's psychological experiences of abortion are shaped by the immediate and larger sociocultural context within which the abortion occurs. From this perspective, social and cultural messages that stigmatize women who have abortions and convey the expectation that women who have abortions will feel bad may themselves engender negative psychological experiences. In contrast, social and cultural messages that normalize the abortion experience and convey expectations of resilience may have the opposite effect.

The co-occurring risk perspective emphasizes that pre-existing and/or ongoing conditions may account for differences in mental health or problem behaviors observed between women who have had an abortion and women who have not. Unwanted pregnancy and abortion are correlated with preexisting and/or ongoing conditions (e.g., poverty), life circumstances (e.g., exposure to violence), problem behaviors (e.g., drug use), and personality characteristics (e.g., avoidance style of coping with negative emotion) that can have profound and long-lasting negative effects on mental health. These conditions may predispose women to unintended pregnancies and abortion and have negative effects on mental health regardless of reproductive history and outcomes. From this perspective then, mental health and problem behaviors observed after abortion are often a byproduct of conditions and characteristics that preceded or coexist with the unintended pregnancy and abortion.

METHODOLOGICAL ISSUES IN ABORTION RESEARCH

Many scholars have noted that research on the mental health implications of abortion is plagued by numerous methodological problems (see, e.g., Adler et al.,

1992; Koop, 1989; Wilmoth et al., 1992). These problems continued to be reflected in most of the studies reviewed by the current task force and limited conclusions that could be drawn from this literature. In the following discussion, we highlight the problems that we encountered most often in our review of the post-1989 literature. We do not recapitulate all of the details presented in previous methodological discussions (see McCall & Appelbaum, 1991, for further discussion of some of these issues). The primary issues we address are those of comparison and contrast groups, co-occurrence of risk factors, sampling, measurement of reproductive history and underreporting, attrition, statistical treatment of data, outcome measurement, and clinical relevance. These issues are not independent of each other. Indeed, the complex interactions among these factors can make it difficult to sort out their separate and combined effects.

Comparison/Contrast Groups

In order for empirical research to address the relative risk of elective abortion compared to alternative courses of action that a pregnant woman facing an unwanted pregnancy might take, clearly defined and otherwise equivalent comparison groups are essential. Otherwise, any previously existing group differences associated with the outcome variable may badly bias conclusions. One appropriate comparison group would be women who are denied or unable to obtain an abortion and who, hence, must carry to term an unwanted pregnancy. Other appropriate comparison groups would be women who deliver an unwanted pregnancy and either give the child up for adoption or raise it. By at least partly controlling for the "wantedness" of the pregnancy, such comparisons provide assurance that the women being compared face a similar situation. Unfortunately, very few studies used appropriate comparison groups.

One way researchers attempted to solve this problem was by using covariate adjustments to try to make "nonequivalent" groups "equivalent." The analysis of covariance, however, can be extremely sensitive to violations of its assumptions, and these assumptions are particularly liable to violation when used to try to adjust for initial group differences (see, e.g., Elashoff, 1969). One violation occurs when the covariate(s) are measured after the treatment—a problem characteristic of retrospective studies of abortion, in which the covariates are assessed after the abortion. A second viola-

tion occurs when the relationship between the covariate and the outcome differs across groups. A third violation occurs when the relationship between the covariate and the outcome is nonlinear. Unfortunately, tests of the validity of these assumptions were rarely encountered in the published literature on abortion. Consequently, caution should be exercised in accepting the findings of studies in which initially incomparable groups were compared (adjusted for covariates) without a test of the validity of the assumptions.

Co-Occurring Risk Factors

Unfortunately, very few studies encountered in our review of the literature adequately assessed and controlled for co-occurring risks. As discussed above, there are naturally occurring interrelations among many of the phenomena associated with elective abortion that make it difficult to tease apart the causal chains that might be operating. Elective abortion commonly co-occurs with unwanted or unintended pregnancy, and unwanted/unintended pregnancy is often associated with adverse circumstances and characteristics that may be associated with mental health problems. Because few studies adequately controlled for these co-occurring risks, it is almost impossible from the available literature to distinguish outcomes that flow from abortion *per se* from outcomes that might appear to be associated with abortion, but in actuality have their origins in the unwanted/unintended pregnancy (or some other co-occurring risk), which is more highly represented in the abortion group than in the comparison group. It was particularly difficult to detect these co-occurring conditions and their consequences from secondary data analyses of data sets collected for other purposes because potential confounds that were not of interest in the initial data collection were unlikely to have been adequately assessed.

Sampling

Problems of sampling characterized most of the studies reviewed. Two basic designs in the abortion literature presented sampling problems. The first occurred when convenience samples of women were recruited specifically for the study without concern for the degree to which they represented a definable population, for example, women seeking pregnancy testing at a health clinic (Cohan, Dunkel-Schetter, & Lydon, 1993), women waiting to see their doctor (Williams, 2001), or pregnant teens residing at a maternity home

(Medora, Goldstein, & von der Hellen, 1993). Often the samples were extremely small (< 30; e.g., Cohan et al., 1993). In many cases, little, if anything, was reported about the inclusion rates of the women in either the abortion group or the comparison groups or the context of their situations, information necessary to establish the representativeness and generalizability of the data. Sometimes data were based on volunteer samples of women who responded to mailed questionnaires about their reproductive history (Reardon & Ney, 2000). Such volunteers do not represent an unbiased sample representative of the population as a whole and cannot be used as evidence to establish prevalence rates or normative responses.

The second and equally problematic situation occurred when subsamples were selected for analysis from extant studies that were initially conducted for other purposes. This characterized most of the studies based on secondary analyses of medical records or public survey data sets. Many of the studies with the largest sample sizes that have been used to make claims about the effects of abortion are of this type—e.g., studies based on the National Longitudinal Study of Youth (NLSY) (e.g., Reardon & Cougle, 2002a; Russo & Zierk, 1992), National Survey of Family Growth (NSFG) (e.g., Cougle, Reardon, & Coleman, 2005), or the National Longitudinal Study of Adolescent Health (Coleman, 2006). In these studies, subsets of the complete sample were taken to allow certain comparisons of interest to be made. For example, only women who reported terminating or delivering a first pregnancy might be selected (e.g., Cougle et al., 2003).

There are a number of serious problems with selecting subsamples from the larger data set in this way: (a) The secondary sampling destroys the sampling properties that might have originally characterized the sample, particularly if population-based sampling weights were not properly taken into account. Distorted sampling weights (or non-use of sampling weights) can lead to inaccurate estimations when the results are used to estimate prevalence of mental health problems in the general population following abortion. (b) Sampling on certain characteristics (e.g., first pregnancy; Cougle et al., 2005; Schmiede & Russo, 2005) may affect other characteristics of the sample, thereby compromising generalizability. For example, women who have an abortion on their first pregnancy are more likely to be younger and to be unmarried than women

who have their first abortion on a later pregnancy. (c) In some studies, additional sources of non-equivalence between abortion and comparison groups were created by selecting a first “target” pregnancy occurring in a specified time period of data collection (e.g., the latter 6 months of 1989). This was to create abortion and delivery comparison groups without attention to reproductive history differences between these groups, when reproductive history is a factor affecting retention in the population sampled (e.g., Cougle, Reardon, & Coleman, 2003; Reardon & Coleman, 2006; Reardon & Cougle, 2002a). (d) Serious violation of sampling principles also occurs when differential exclusion is used in constructing comparison groups such that one group is advantaged relative to the other (e.g., Coleman et al., 2002; Cougle, et al., 2005).

Measurement of Reproductive History and Problems of Underreporting

Many of the studies reviewed were characterized by inaccuracy in the information available regarding a woman’s reproductive history, particularly her abortion history. In some studies, a woman’s abortion status was verifiable (e.g., data were collected at the time that she sought an abortion at a clinic or from her medical records). More typically, however, abortion status was established based on self-report. For example, in all of the studies based on a secondary analysis of survey data, abortion status was established by asking women to indicate, either on a questionnaire or verbally, to an interviewer whether or not they had had an abortion in the past. Women’s reports of an earlier abortion were then correlated with current mental health/emotional status, with the latter attributed to the former (e.g., Coleman, Reardon, Rue, & Cougle, 2002a; Cougle et al., 2005).

This approach has many problems. Abortion, like other stigmatized conditions, is typically underreported (Jones & Kost, 2007). It has long been recognized that individuals are unlikely to frankly answer questions that have the potential to be embarrassing, overly self-disclosing, or in other ways reflect negatively on them. One of the earliest applications of a statistical model designed for reducing bias in obtaining answers to sensitive questions—the so-called randomized response methodology—was for estimating the mean number of abortions in an urban population of women (Greenberg, Kuebler, Abernathy, &

Horvitz, 1971). The percentage of women reporting an abortion on surveys is consistently lower than the number expected based on estimates made from national provider data, sometimes markedly so (Jones & Forrest, 1992; Jones & Kost, 2007). Absent the use of techniques such as randomized response methodology or the selection of highly disclosing samples, one is likely to obtain biased estimates of prevalence rates. Generally, there are two types of underreporting: failure to acknowledge having had any abortions and having had multiple abortions but reporting only some of them (Jones & Kost, 2007).

Underreporting of abortion in surveys is of particular concern when there is differential underreporting by subgroups of women (Fu, Darroch, Henshaw, & Kolb, 1998; Jones & Forrest, 1992). Women more likely to underreport include those who are unmarried, Black or Hispanic, Catholic, low-income, and aged 20–24 (Jones & Kost, 2007). Underreporting can introduce systematic bias into a study. Only a few studies reviewed attempted to test for possible underreporting biases. For example, Schmiede and Russo (2005) examined and compared the relation of abortion versus delivery to depression (CESD cutoff score) in the NLSY data set among groups known to vary in underreporting (e.g., White married women, unmarried Black women, Catholics). Their analyses suggested that at least in the NLSY data set, underreporting by specific subgroups did not appear to introduce systematic bias into observed associations between abortion and a mental health outcome.

In general, the nature of the potential bias introduced by underreporting (i.e., whether it biases toward overestimating or underestimating adverse impact of abortion) is unclear. It is possible that women who feel most distressed by an abortion are less likely to report it to others; as a consequence, they may be underrepresented in the abortion group, biasing results toward underestimating negative effects. It is also possible that response biases in the other direction may be observed. For example, women who are experiencing distress may view the survey as an opportunity for catharsis and hence be more likely to disclose their abortion than women less distressed. In addition, women most willing to report one “problem” (e.g., depression, anxiety, abuse) may also be those most able to recall or willing to report another “problem behavior” (abortion), biasing results toward overesti-

mating negative effects. Many scholars have noted the problem of selective recall bias in surveys on the part of individuals experiencing a disorder who may (1) more thoroughly scrutinize their history in an effort to explain their disorder and (2) more accurately recall stigmatizing events, such as abortion, than individuals not experiencing a disorder (e.g., Neugebauer & Ng, 1990; Chouinard & Walter, 1994). Recall biases can explain, for example, why a positive relationship between abortion history and breast cancer has been observed in retrospective surveys but is absent in prospective studies (American Cancer Society: <http://www.cancer.org/>). Specifically, breast cancer patients seeking to understand their disease are thought to be more motivated to search their memories as well as more willing to report socially stigmatizing conditions (such as abortions or sexually transmitted infections) to a health care provider than are healthy women, leading to a spurious relationship.

Measurement of abortion also typically suffers from underspecification. Many studies lack important information about the abortion, such as length of gestation, type of procedure, or whether the abortion was performed for therapeutic reasons, all of which may affect how women respond emotionally and physically after an abortion. For example, abortions performed beyond the first trimester involve a more risky medical procedure and more pain, which may have negative effects. They also occur at a more advanced stage of development, which may change the meaning of the pregnancy, making abortion more stressful (Major, Mueller, & Hildebrandt, 1985). Delay may also reflect ambivalence toward the pregnancy or indicate that a wanted pregnancy was terminated because of discovery of a health problem or fetal defect. It is also unclear to what extent research on earlier surgical methods of abortion applies to newer nonsurgical methods of abortion, which are used at the earliest stages of gestation and differ from traditional methods in other ways as well, although studies suggest comparable postabortion emotional adjustment for women experiencing each method (Ashok, Hamoda, Flett, Kidd, Fitzmaurice, & Templeton, 2005; Howie, Henshaw, Najo, Russell, & Templeton, 1997; Lowenstein et al., 2006; Sit, Rothschild, Creinin, Hanusa, & Wisher, 2007).

Attrition

Another potentially serious methodological confound encountered was attrition—loss of cases during the

course of an investigation. Attrition has been a long-standing concern in the study of abortion (see for example, Adler, 1976). The consequences of attrition range from potentially serious loss of power to biasing of results when attrition is not random (i.e., biased in a specific direction) and differs by group. In the case of abortion, for example, underestimation of the prevalence of distress in the final sample would occur if women who were most upset by the abortion were more likely to be lost to a follow-up than those who were retained in the sample. Similarly, overestimation of the prevalence of distress would occur if women who were least distressed by the abortion were more likely to be lost to a follow-up. Consequently, it is essential that researchers test for biases in attrition. Only a few studies reviewed did so. One study that did test for attrition (Major et al., 2000) found that among women who had a first-trimester abortion, those who were retained in the sample at the 2-year postabortion measurement period did not differ significantly from those who were lost to attrition on any demographic or psychological variable assessed either prior to the abortion, immediately post abortion, or 3 months post abortion. Thus, at least in this sample, no evidence of systematic bias in attrition was observed.

Outcome Measures: Timing, Source, and Clinical Significance

Problems of outcome measurement also were frequently encountered in this literature. It is vital that the measures of mental health are valid and reliable. In some studies reviewed, claims of mental health impact (or no impact) were made on the basis of psychometrically poor measures, including one-item measures (e.g., Coleman, 2006a; Reardon & Ney, 2000). For example, Reardon and Ney (2000) measured substance abuse with yes/no responses to the single question “Have you ever abused drugs or alcohol?” This is not a reliable measure of substance abuse. A clinically relevant measure (as opposed to a scale score without known clinical relevance) should be the minimal standard for measuring impact. In addition, claims of impact should be accompanied by epidemiologically meaningful effect size indicators such as odds ratios, which provide clinically relevant measures of impact. Odds ratios should be presented in conjunction with data of the rates or proportions of women affected (i.e., a finding of 3 to 1 in 100 women presents a different level of threat than 3 to 1 in 1 million women). Absolute and relative levels of the effect should be clear.

An associated problem encountered in both primary and secondary studies was related to the timing of measurement. Some studies first contacted their participants months or years (or an unspecified time interval) after the target abortion and engaged them in retrospective reporting of their preabortion status (e.g., Bradshaw & Slade, 2005; Cougle et al., 2005) or their mental health/emotional status at selected points after the event (e.g., Kersting et al., 2005). Retrospective reporting is subject to a large number of distortions and biases. There is agreement among methodologists that measures taken nearer an event are more likely to be accurate than measures taken at a time distant from the event.

Finally, assessing the clinical significance of abortion, as with any other medical procedure, requires asking “what is the benefit?” as well as “what is the harm?” of the procedure. Many of the abortion studies reviewed focused only on negative outcomes. Focusing solely on adverse effects can create a distorted picture of the information needed to provide complete and accurate informed consent. It is akin to focusing on the risks of chemotherapy without addressing its potential benefits for curing cancer. For example, in separate reports based on the same sample, one research team reported a negative association between abortion and mental health (Fergusson, Horwood, & Ridder, 2006) and a positive association between abortion and other life outcomes (e.g., education, employment; Fergusson, Boden, & Horwood, 2007). The authors concluded that there is a “need for further research into the risks and benefits associated with abortion as a means of addressing the issues raised by unwanted or mistimed pregnancies” (Fergusson et al., 2007, p. 11).

Other Statistical Issues

Many of the studies included in our review were characterized by statistical problems. One frequently encountered problem, especially in the studies based on secondary data analyses, was inflation of the probability of making a Type I error in inference by performing many significance tests at the same level one would if there were to be only a single test. This appeared in two forms. The first form occurred when the initial sample (often a reasonably large sample) was divided into smaller and smaller subsets, and these subsets were then used to test for differences between abortion and nonabortion cases within each subset without any overall control for the number of signifi-

cance tests conducted (e.g., Coleman, Reardon, & Cougle, 2002; Reardon & Ney, 2000). This practice increases the probability of a statistically significant difference occurring due to chance. The second form encountered was the ad hoc search for covariates. In many studies, especially those based on analyses of secondary data sets, the data analyst began with a set of all possible covariates (usually defined by the measures available in the data set) and tested each covariate for significance (testing the partial regression coefficients for significance). The analyst then proceeded to conduct analyses using only the significant covariates (e.g., Coleman, Maxey, Rue, & Coyle, 2005). Without any correction for chance via alpha-level control, this completely ad hoc, atheoretical approach also capitalizes on chance. Furthermore, the choice of covariates to include in analyses can play a key role in how much variance in the outcome variable is explained by pregnancy outcome.

Interpretational Problems and Logical Fallacies

In addition to the methodological problems described above, the TFMHA also encountered a number of cases in which data were incorrectly interpreted or generalized, if not in the actual research reports themselves, then in reviews, summaries, and press releases based on that research. Accordingly, the TFMHA felt it important to point out several logical fallacies that must be guarded against in drawing conclusions from this literature.

The first logical fallacy is the tendency to infer causation from correlation. Frequently, significant correlations observed between abortion history and other variables (e.g., substance abuse, depression, higher educational outcomes) were misinterpreted as evidence that abortion caused these variables to occur. Such causal claims are unwarranted, as the relationships may be spurious, the causal direction may be reversed, or the relationship may be due to a third variable that is associated with both abortion and the outcome variable (e.g., poverty). It is sometimes argued that a case for causality is stronger in abortion studies that establish (a) time precedence of the abortion before an outcome variable, (b) covariation of abortion and the outcome variable, and (c) lack of plausible alternative explanations or control of third variables associated with both abortion and the outcome variable. These, however, are only necessary but not sufficient conditions to establish causality. Furthermore, although

some of the studies reviewed did meet criteria (a) and (b), the TFMHA could identify no study reporting a significant association of abortion with a mental health outcome that met criterion (c).

A second logical fallacy is the tendency to confuse a risk and a cause. For example, some writers appeared to assume that if a prior history of abortion was found to be a “risk factor” for a certain outcome (e.g., violent death), then a prior history of abortion is a “cause” of violent death. Many things can serve as markers for causes or may be associated with causes without themselves being a part of the causal mechanisms in play. For example, age is the most important known risk factor for Alzheimer’s disease (AD), but it is not the mechanism that causes people to develop AD. Rather, age is a statistical predictor in a population of who in that population is at risk, that is, more likely (older versus younger) to develop AD (<http://www.nia.nih.gov/Alzheimers/>). The steps that link risks and causes must be explicitly developed and demonstrated before one can validly make the assertion that removing a particular risk factor will lead to a desired outcome.

A third and very serious logical fallacy is the “interventionist fallacy”—the belief that if a relationship is observed between two variables, the form or magnitude of the relationship will remain unchanged if an intervention changes some part of the current state of affairs. For example, because there is a substantial positive relationship between family income and children’s school performance, it is tempting to think that increasing family income would lead to improved children’s school performance. Such a conclusion, however, does not logically follow. It might be that what drives the relationship between family income and school performance is the family expenditure on books. Were one to intervene and supplement family income, it does not necessarily follow that the family would increase its expenditure on books, which are (in this example) the actual component that drives the child’s school performance, and, hence, the intervention might fail.

As applied to the case of abortion, one example of the interventionist fallacy would be the belief that if abortion and depression are related, then reducing access to abortion would reduce the prevalence of depression. A change in the availability of elective

abortion, however, would have many consequences. It would mean that women who want to terminate an unwanted pregnancy would now be forced to deliver. As a consequence, the characteristics of the population of women who delivered children would change. Characteristics previously prevalent among women who had an abortion (e.g., greater poverty, exposure to violence) would now be prevalent among the delivery group. The portrait of the mental health of mothers might reasonably be expected to be worse. This potential change in the profile of women giving birth does not include any new mental health problems that might develop from stresses associated with raising a child a woman feels unable to care for, or may not want, or from relinquishing a child for adoption. Thus, reducing access to abortion would be likely to result in poorer mental health among women who deliver. Hence, rather than reducing the prevalence of depression among women, this intervention could potentially increase it.

Summary of Methodological Issues

Most of the studies published on postabortion mental health contain one or more of the methodological or interpretational problems discussed above. Consequently, reviews of the literature that simply count the number of studies that show one effect versus another or that calculate effect sizes without carefully considering and weighing the quality of the evidence that produced the effect are inappropriate and often misleading. It is essential to keep the methodological and interpretational points discussed above in mind when considering the literature on postabortion mental health reviewed below.

It is also important to recognize, however, that not all design problems are equally serious. The extent to which a design flaw affects the merits of a particular study depends in part on the goal of the study. For example, the lack of a comparison group is not overly limiting when the researcher’s goal is to understand predictors of response among women who have abortions. Some flaws can be compensated for by limiting generalization or interpretation. However, other flaws are so serious that they limit any conclusions that can be drawn from the study (e.g., differential exclusion of women from one group but not the comparison group on a variable known to be related to the outcome variable).

REVIEW OF SCIENTIFIC LITERATURE

Search Strategy and Criteria for Inclusion

In order to evaluate the scientific literature on mental health effects of abortion, the TFMHA searched PsycINFO and Medline for English-language peer-reviewed articles published between 1990 and 2007 based on human subjects. Research conducted with U.S. as well as non-U.S. samples was searched. Key-word combinations paired *abortion* with each of the following words: *anxiety, depression, mental disorders, mental health, trauma, PTSD, domestic violence, drug abuse, emotions, employment, life satisfaction, self-esteem, somatoform, stigma, substance abuse, suicide, acute psychosis, schizophrenia, psychiatric symptoms, and psychosocial factors*. In addition, *postabortion syndrome, postabortion adjustment, and therapeutic abortion* were also used as search terms. The search results were supplemented by a manual search of reference sections of reviewed articles. This search strategy resulted in an initial set of 216 unique references. Seven additional references were brought to the attention of the task force by reviewers.

Our review process consisted of four steps. In the first step of review, the abstract of each article in the initial set was reviewed independently by two task force members according to the following inclusion criteria: (1) The study reported empirical data of a quantitative nature (qualitative studies were omitted). (2) The study was published in a peer-reviewed journal (dissertations, letters to editors, reviews, book chapters, and conference proceedings were omitted). (3) The study included at least one postabortion measure related to mental health (those that considered only mental health prior to the abortion were omitted). (4) The study focused on induced abortion [those that focused solely on "spontaneous" abortions (miscarriages) or that did not differentiate miscarriage from induced abortion were omitted].

Those articles that appeared to meet all of the above criteria were included for further review. In the second step, a minimum of two task force members independently read all articles identified in our first step. Only articles judged to have met all of the above inclusion criteria were retained. In the third step, all studies that met criteria for inclusion were coded, summarized,

and evaluated independently by at least two members of the task force, with the restriction that task force members did not evaluate their own work.

In a final step, articles were categorized according to whether or not they included a comparison group of women who did not have an abortion. Only studies that include a comparison group are capable of addressing the question of relative risk. Accordingly, our core review focused only on studies that included comparison groups. Studies without a comparison group have the potential to address predictors of individual variation in women's responses following abortion. They also are capable of addressing the question of prevalence of mental health problems among women who have abortions, but only to the extent that they are based on a sample representative of the population to which one intends to generalize. Accordingly, in a separate section we review such studies, but only when based on a U.S. sample.

Descriptive Overview of Literature Identified for This Review

Through the process described above, 50 papers were identified that compared psychological experiences of women after abortion to psychological experiences of a comparison group of women. These 50 include studies based on U.S. and international samples. The restriction of empirical studies to those published in English resulted in a relatively narrow slice of international contexts represented in this report. One should not assume that this small set is representative of the global experience of abortion and mental health, as laws, customs, and contexts vary widely. Twenty-five papers compared women who had an abortion to women who had a different reproductive history (e.g., a delivery, miscarriage, no pregnancy) by performing secondary analyses of public data sets or records originally collected for other purposes; 18 of these papers were based on U.S. samples; the remaining papers were based on samples from New Zealand (1) and Finland (6). These are summarized in Tables 1 and 2. A second set of papers ($N=19$) described original studies conducted primarily for the purpose of comparing responses of women who had a first-trimester abortion (or an abortion of unspecified gestation) to responses of women who had a different reproductive history. Most of these studies were based on samples collected at clinics or physicians' offices; some were retrospective. Seven were conducted in the United States, the

remainder in other countries. These studies are summarized in Tables 3a and 3b. A third set of papers ($N=6$) consisted of studies comparing psychological experiences of women who had a late-trimester abortion of a pregnancy for reasons of fetal anomaly to another group of women. All but one was conducted on non-U.S. samples. These studies are summarized in Table 4. These 50 papers constitute the core of our review. Our literature search also identified 23 papers based on U.S. samples that did not include a comparison group but met all other inclusion criteria. These papers are summarized in Table 5.

REVIEW OF COMPARISON GROUP STUDIES

Record-Based Studies and Secondary Analyses With Comparison Groups

The major change in the scientific literature during the time period encompassed by our review compared to the literature reviewed by the first APA task force was the publication of 25 papers in peer-reviewed journals based on secondary analyses of publicly available data sets. The studies are of two types: (a) analyses of data based on medical records and (b) analyses of data sets collected for purposes other than analyzing the relationship between pregnancy outcome and mental health. Because publicly available data sets often include questions about reproductive histories, including pregnancy outcomes (abortion, delivery, miscarriage), they provide an opportunity for comparing women who report having had an abortion to other groups of women. Utilizing existing data sets, particularly longitudinal data sets, also has the advantage of being able to ask and answer questions without having to wait the years it takes to conduct a prospective study focused specifically on abortion. Findings based on national probability samples potentially may be generalized more widely than those based on convenience samples and may be more useful for estimating normative effects. Nonetheless, as pointed out above in the methodological issues section of this report, there are many serious limitations of this approach that severely constrain conclusions that can be drawn from these studies (see also McCall & Appelbaum, 1991). In the following discussion, we provide a brief description of these studies, followed by an evaluation of

their methodology. Table 1 and Table 2 provide a description of the key methods, measures, and findings of these studies, as well as their limitations.

Medical records. Ten papers were published based on medical records. Four papers were based on analyses of medical records from California's state-funded insurance program (Medi-Cal). This program provides health care for low-income children and families, as well as elderly, blind, and disabled persons in the state of California. These "at-risk" women may be facing a wide range of challenges that compromise their physical and mental health. Six reports were based on official health register data drawn from medical records and on the entire population of Finland (See Table 1).

All four Medi-Cal studies focused on an initial target pregnancy event (abortion vs. delivery) in the last half of 1989 and after excluding women with subsequent abortions only from the delivery group, examined the records of the remaining sample of women for subsequent death (Reardon et al., 2002), outpatient admissions (Coleman, Reardon, Rue, & Cougle, 2002b), inpatient admissions (Reardon, Cougle, Rue, Shuping, Coleman, & Ney, 2003), and sleep disturbances (Reardon & Coleman, 2006). All four papers reported higher rates of negative outcomes in the abortion group compared with the delivery group.

In considering the weight of the evidence with regard to the mental health implications of abortion, it should be kept in mind that the Medi-Cal studies are not independent of each other because the samples overlap, and most of the outcomes examined are correlated. Strengths of the Medi-Cal studies include an objectively verifiable abortion history and the use of diagnostic codes for assessing mental illness. Nonetheless, these papers are characterized by a number of methodological limitations that make it difficult to interpret the results. These include differential exclusion of women with subsequent abortions from the delivery group but not from the abortion group, a sampling strategy that both advantaged the delivery group and rendered generalizability of the findings problematic; lack of basic demographic information known to be associated with mental health, including marital status and race; lack of information about previous reproductive history, lack of adequate assessment of prior mental health history, lack of adequate information about co-occurring risks (e.g., health status, vio-

Table 1: Medical Records—Abortion vs. Comparison Groups
U.S. STUDIES

Medi-Cal Data Set

General Description: Medi-Cal is California's state-funded medical insurance program for low-income individuals; 249,625 women identified as having a "short paid claim" for Medi-Cal funding for either an abortion or delivery (pregnancy event) in 1989; for most studies 194,694 women were identified as citizens with valid SSN. Samples for the studies below were based on this subgroup. In 1989, pregnant women were Medi-Cal-eligible if family income was less than 185% of federal poverty level.

Limitations Common to All Studies Based on this Data Set: Pg intendedness or wantedness not controlled; basic covariate info (e.g., race, marital status, # births & abortions) unknown; inadequate controls for prior mental illness; sample representativeness suspect, even for generalizing to low-income population—more than 20% of the sample excluded before samples of specific studies selected. Pg outcome may affect eligibility in different ways: having a baby may qualify a woman for Medi-Cal, independent of her own characteristics, while women who remain on Medi-Cal postabortion would have to qualify for other reasons.

Causality direction ambiguous—women with poor health status may be more likely to choose abortion. Misleading "First Pregnancy" label used to identify target population cannot be specified; N is so large that minute differences can be statistically sig. Impact of controlling months of eligibility not clear as women may have lapses of coverage during period examined. Given poorer health of low income populations, inability to separate therapeutic from elective abortions a particular limitation.

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Additional Limitations Specific to Study Listed
Reardon, D.C., Ney, P.G., Scheuren, F., Cogle, J., Coleman, P.K., & Strahan, T.W. (2002). Deaths associated with pregnancy outcome: A record linkage study of low income women. (2002) <i>Southern Medical Journal</i> , 95, 834-841.	Medi-Cal records linked to death certificates between 1989-89; after screening for "aberrant, indeterminate, and out-of-scope data" 173,279 (1,294 deaths) cases used in primary analyses; some analyses excluded women with inpatient and outpatient psychiatric claims in preceding 6-18 months & women with subsequent abortion from delivery group.	1. AB N= 30,260 DEL N= 83,690 2. AB N= 41,956 DEL N= 17,472	Rates of causes of death reported on California death certificate between 1989 and 1998.	1. Age-adjusted risk of death significantly higher in AB group from violent causes but not for nonviolent causes. 2. Women with subsequent abortions were excluded only from DEL group and number of psychiatric claims in previous yr controlled, age-adjusted risk of death significantly higher in AB group for both violent and nonviolent causes.	39,329 (65%) of excluded sample is without explanation; 8-year period in which deaths identified not congruent with ACOG definitions of pregnancy-related or even pregnancy-associated death. Differential exclusion advantages delivery group.
Coleman, P. K., Reardon D.C., & Rue V.M. (2002). State-funded abortions versus deliveries: A comparison of outpatient mental health claims over 4 years. <i>American Journal of Orthopsychiatry</i> , 72, 141-152.	Differs from general description above in reporting 193,794 women as having a valid SSN; after screening exclusions, women with target Pg event in the last half of 1989 selected; women with subsequent abortion excluded from delivery group; women with both inpatient or outpatient psychiatric admissions claim in yr preceding target Pg event excluded; final sample = 54,419.	AB N= 14,297 DEL N= 40,122	1. Cumulative rates of outpatient psychiatric admission claims at 90d, 180d, yr1, & yrs 1-4 after target Pg event; 2. Rates of disorder in 13 groups of selected ICD-9 diagnostic categories.	1. Significantly higher cumulative rates of outpatient claims for AB group controlling for age, number of Pg events, & months of Medi-Cal eligibility. 2. Of 13 comparisons, AB group rates significantly higher in 4 categories (adjustment reactions; bipolar disorder; neurotic depression; schizophrenic disorders); marginally significant in 2 (anxiety states; alcohol & drug abuse).	Differential exclusion of women with subsequent abortion from DEL grp; inadequate control through exclusion for prior mental disorder; flaw of "validation by cross-quotation" found in claim of evidence for causal model "accumulating" based on citation to research that does not warrant that claim (p. 149).

Table 1: Medical Records–Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Additional Limitations Specific to Study Listed
Reardon, D.C., Cogle, J. R., Rue V. M., Shuping M. W., Coleman P.K., & Ney P.G. (2003). Psychiatric admissions of low-income women following abortion and childbirth. <i>Canadian Medical Association Journal</i> , 168, 1253-1256.	After screening exclusions, women with target Pg event in the last half of 1989 were selected; women with inpatient psychiatric admissions claim in year preceding target Pg event excluded; women with subsequent abortion excluded from delivery group; final sample 56,741.	AB N= 15,299 DEL N= 41,442	1. Cumulative rates of inpatient psychiatric admission claims at 90d, 180d, & yr1 after target Pg event; 1st time rates in yr 1, 2, 3, & 4 after target Pg event; 2. Rates of disorder in 9 groups of selected ICD-9 diagnostic categories.	1. Controlling for age and months of Medi-Cal eligibility to the end of the time period analyzed, the AB group had significantly higher rates for both cumulative and 1st time rates of inpatient claims for AB group at time periods listed; 2. Of 9 comparisons, rates of AB group were significantly higher in 4 categories (adjustment reaction; depressive psychosis; single episode; depressive psychosis, recurrent episode; bipolar disorder).	Reluctance to hospitalize new mothers could account for lower post-delivery admission rates. Misleading use of term "first admission" because only mental health claims for one year prior to Pg were examined. Inadequate controls for prior mental illness.
Reardon, D.C., & Coleman, P.K. (2006). Relative treatment for sleep disorders following abortion and child delivery: A prospective record-based study. <i>Sleep</i> , 29, 105-106.	After screening exclusions, women with a history of treatment for sleep disorder excluded; women with subsequent abortion excluded only from delivery group; final sample = 56,824 cases.	AB N= 15,345 DEL N= 41,479	Cumulative rates of treatment for category representing nonorganic sleep disorder and sleep disturbances at 180d, yr1, and 1-4 years after target Pg event; 1st time rates yr 1 through 4 after target Pg event.	Controlling for age and number of months of Medi-Cal eligibility, significantly higher treatment rates in AB group at 180 d, yr1 & yr 4, & significantly higher 1st time rates in yr 3, but not yrs 2 & 4.	Impact of controlling for months of eligibility is not clear as authors note that some women had lapses of coverage during the period examined.

lence exposure), lack of information about critical characteristics of the abortion decision context (e.g., whether the pregnancy was initially intended and terminated because of fetal anomalies), and inclusion of covariates across analyses and studies that varied for unspecified reasons (see Table 1). Yet another problem with this data set is that women who deliver a child are more likely to be eligible for Medi-Cal because they have a baby, independent of their own characteristics. Women who have an abortion may qualify for the abortion, but those who remain on Medi-Cal post abortion (and who hence would be picked up in the follow-up measurement) would have to have other characteristics besides motherhood to

qualify (e.g., mental illness, other illness, poverty not associated with parenthood).

The Medi-Cal findings with regard to cause of death (Reardon et al., 2002) can be compared with record-based studies conducted in Finland that are based on the entire population of the nation (Gissler, Hemminki, & Lonnqvist, 1996; Gissler et al., 1997), albeit from a differing cultural context. These studies also found significantly higher rates of pregnancy-associated deaths for natural and violent causes (including accidents, homicide and suicide) in the abortion group compared with a delivery group. Like the Medi-Cal studies, these studies also had

Table 1: Medical Records—Abortion vs. Comparison Groups
INTERNATIONAL STUDIES—Finland

General Description: National data registers based on medical records make it possible to examine health status of the entire population of the country so underreporting bias not a major issue. These studies provided inspiration for Medi-Cal studies. Note outcomes based on ACOG definitions of Pg-associated deaths (occurring within one year of end of pregnancy, regardless of cause of death) vs. Pg-related deaths (occurring within one year of end of Pg from any causes related to or aggravated by their Pg or its management, but not from accidental or incidental causes) differ from definitions in Medi-Cal studies.

Limitations Common to All Studies Based on this Data Set: Neither intendedness nor wantedness of Pg controlled; information on age, marital status, and reproductive history lacking; low rates of unintended pregnancy and ready access to abortion in Finland make it likely most births are wanted.

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Additional Limitations Specific to Study Listed
Gissler, M., Hemminki, E., & Lonnqvist, J. (1996). Suicides, 1987-94: register linkage study. <i>British Medical Journal</i> 313, 1431-1434.	Death register records for 1347 suicides were linked to birth, abortion, and hospital discharge records, identifying 73 deaths occurring within 1 year of a birth or abortion.	AB N= 29 DEL N= 30	Suicide rates	Suicide rate significantly higher in AB group: Divorced women and women in the lower social classes were over-represented in the AB suicide group vs. women in the abortion register overall.	Given findings on class and marital status in AB group, lack of control for wantedness, exposure to violence, class, parity, and circumstances of the Pg makes comparisons between AB and DEL groups problematic.
Gissler, M., Kauppila, R., Merilainen, J., Toukomaa, H., & Hemminki, E. (1997). Pregnancy-associated deaths in Finland 1987-1994—Definition problems and benefits of record linkage. <i>Acta Obstet Gynecol Scand</i> , 76, 651-637.	Record linkage study of women of reproductive age between 1987-1994; 281 deaths identified as Pg-associated.	AB N= 84 Miscarriage N= 40 DEL N= 137	Rates of causes of death	Higher age-adjusted rates for overall deaths, natural deaths, accidents, suicides, & homicides in AB group.	Pg-related deaths not identified. Only age controlled.
Gissler, M., & Hemminki, E. (1999). Pregnancy-related violent deaths. <i>Scand J Public Health</i> , 1, 54-55. [Letter to editor].	Additional analyses of violent death. Identified record linkage study of violent deaths among the 281 Pg-associated deaths identified in Gissler et al (1997).	AB N= 84 Miscarriage N= 40 DEL N= 138	Rates of causes of violent death	Higher age-adjusted rates of accidents, suicides, & homicides in AB group.	Only age controlled. These data are based on the same records as Gissler et al. (1997) & apparently were an attempt to counter claims that Gissler et al (1996) implied causation. Authors emphasize the point that given the "finding that the risks for accidental death and homicide also increase after an induced abortion and our previous findings that women from lower social classes and single women are over-represented among women committing suicides after an induced abortion, do not support the hypothesis that abortion itself causes suicides" (p. 55).

Table 1: Medical Records—Abortion vs. Comparison Groups
INTERNATIONAL STUDIES—Finland (continued)

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Additional Limitations Specific to Study Listed
Gissler, M., Berg C., Bouvier-Colle, M. H., & Buekens, P. (2004). Methods to or identifying pregnancy-associated death: population-based data from Finland 1987-2000. <i>Pediatric and Perinatal Epidemiology</i> , 18, 448-455.	Record linkage study of women of reproductive age between 1987-2000; 419 deaths identified as Pg-associated.	AB N= 129 DEL N= 224	Rates of causes of Pg-associated deaths	Higher Pg-associated mortality rates for abortion compared to birth	Pg-related deaths not identified; nothing was controlled.
Gissler, M., Berg C., Bouvier-Colle M. H., Buekens P. (2004). Pregnancy-associated mortality after birth, spontaneous abortion, or induced abortion in Finland, 1980-2000. <i>American Journal of Obstetrics and Gynecology</i> , 190, 422-427.	Record linkage study of Pg-associated deaths 1987-2000; of the 15,823 women who died, 419 of the deaths were Pg-associated.	AB N= 129 DEL N= 224	Rates of natural and violent causes of Pg-associated and Pg-related deaths	1. Pg-associated death rates from natural causes (particularly natural causes unrelated to Pg) & from violent causes higher in AB group. Direct Pg-related causes higher in DEL group, but significance not reported (3.9 & 1.3/100,000 Pg). 2. When therapeutic abortions excluded, Pg-associated mortality rates higher in the DEL group.	Only age controlled. These findings include 1987-2000 cases used in previous studies, so are not independent. Therapeutic abortions in early Pg likely under identified.
Gissler, M., Berg C., Bouvier-Colle M. H., & Buekens, P. (2005). Injury deaths, suicides, and homicides associated with pregnancy, Finland, 1987-2000. <i>European Journal of Public Health</i> , 15, 459-463.	Record linkage study of Pg-associated deaths 1987-2000 from external causes; of the 5,299 women who died, 212 of the deaths were Pg-associated.	AB N= 92 DEL N= 81	Pg-associated deaths from external causes	2. Death rates higher in AB group then DEL group for all external causes, including rates for unintentional injuries, suicide, & homicide.	These findings include 1987-2000 cases used in previous studies, so are not independent. Therapeutic abortions in early Pg likely under identified. Only age controlled. Authors state that their findings do not warrant causal conclusions and emphasize the need for more information on relevant covariates, including "mental health, social well-being, substance abuse, and socio-economic circumstances" in further analyses (p. 462.)

Notes: AB = Abortion; DEL = Delivery; Pg = pregnancy; ACOG = American College of Obstetricians and Gynecologists; ICD = International Classification of Diseases; Grp = Group; Sig = Significance

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES

National Longitudinal Survey of Youth (NLSY)

General Description: The National Longitudinal Survey of Youth (NLSY) is based on annual interviews with a stratified, multi-stage national probability sample of noninstitutionalized civilian men and women aged 14-21 as of 1979, with oversampling of Blacks, Hispanics, and poor Whites. Relevant measures include: an abbreviated version of the Rotter internal-external locus of control scale (Rotter, 1966; IE assessed in 1979); global self esteem (Rosenberg, 1979; RSE assessed in 1980 & 1987); Center for Epidemiological Studies-Depression Scale (Radloff, 1977; CESD assessed in 1992); reproductive histories were first taken in 1982 and updated every 2 years subsequently.

Limitations Common to All Studies Based on this Data Set: No study used sampling weights so that normative statements are inappropriate and alpha levels are likely to be elevated, increasing probability of identifying difference due to chance as a reliable difference. Underreporting of abortion raises question of possible reporting bias but direction of reporting bias unclear as women may be less likely to report stigmatized experiences (having an abortion, mental problems, experiencing violence), but those who are willing to report one stigmatized condition may be more willing to report others, increasing the likelihood of finding a correlation between 2 stigmatized events. Ns of analyses vary depending on covariates so are not always clear. Large sample sizes mean that small effects are statistically significant.

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcomes	Key Findings	Notes and Additional Limita- tions Specific to Study Listed:
Russo, N.F., & Zierk, K.L. (1992). Abortion, childbearing, and women's well-being. <i>Professional Psychology: Research and Practice</i> , 23, 269-280.	1. 5,295 women for whom there were NLSY interviews involving the assessment of well-being in 1987; 773 had at least one abortion; 233 had repeat abortions. 2. Additional analyses based on 4502 women who had no abortions before their 1980 interview.	1. AB N = 733 Other N = 4562 2. AB N = 317 Other N = 4185	1987 Global self-esteem (RSE)	1. Women who had 1 abortion had higher SE than other two groups; when childbearing and resource variables were controlled, neither having 1 abortion nor having repeat abortions were significantly related to RSE. Total abortions correlated with total unwanted births ($r=.11$). 2. 1980 RSE was the strongest predictor of 1987 SE (partial $r=.38$).	No clinical cut off score & clinical significance of scores is unknown; large sample means small effects statistically significant. Limited to women under 33 years of age in 1987.
Russo, N.F. & Dabul, A.J. (1997). The relationship of abortion to well-being. Do race and religion make a difference? <i>Professional Psychology: Research and Practice</i> , 28, 23-31.	1. 4,913 women drawn from the sample of 5,295 women described above (3,572 White & 1,341 Black); 721 had at least one abortion, 175 had repeat abortions. 2. Additional analyses based on 4,336 women (3,147 White & 1,189 Black) who had no abortions prior to 1980 interview.	1. AB N = 721 Other N = 4,192 2. AB N = 317 Other N = 4,502	1987 Global self-esteem (RSE)	Primary findings did not vary across groups known to vary in under-reporting. 1. When childbearing and resource variables were controlled, neither having 1 abortion nor having repeat abortions significantly related to RSE, regardless of race or religion. 2. 1980 SE was the strongest predictor of 1987 SE (partial $r=.39-.42$) regardless of race or religion.	Religion measured in 1979 only; highly committed fundamentalist women not identified; sample does not include Asians or Native Americans. Limited to women under 33 years of age in 1987.

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Notes and Additional Limitations Specific to Study Listed:
Reardon, D.C., & Cougle, J.R. (2002). Depression and unintended PG in the National Longitudinal Survey of Youth: A cohort study. <i>British Medical Journal</i> , 324, 151-152.	Two samples were drawn due to coding issues in the initial study; both the initial and corrected sample ns are reported here. 1. Initial sample: 421 women identified as reporting a first unintended Pg between 1980 and 1992 that resulted in abortion (N=293) or delivery with no subsequent history of abortion in the delivery grp (N=128). 2. Corrected sample: 1076 women identified as reporting a first unintended Pg between 1980 and 1982 that resulted in abortion (N=293) or delivery with no subsequent history of abortion in the delivery grp (N=783). Results were similar in both samples & only results of corrected sample presented here.	1. AB N=293 DEL N=128 2. AB N=293 DEL N=783	% women exceeding the 1992 CESD cut-off score (>15).	AB grp had higher % scoring ≥ 16 on CES-D in 1992 (27% vs. 25%), controlling for family income, education, race, age at 1st Pg, and 1979 I-E score. Significantly higher risk for AB grp among married women (26% vs. 19%), but not among unmarried women (36% vs. 29%, ns), controlling for family income, education, race, age at 1st Pg, and 1979 I-E score.	Note: Differs from RSE studies in focusing on outcome of 1st Pg. Subsequent reanalysis by Schmiege & Russo (2005) showed that findings in corrected sample still based on miscoded data. Exclusion of women with subsequent history of abortion from the delivery group. Uses I-E score as a control for pre-existing mental health but scale is not a measure of mental health. CESD controversial due to cutoff at >15 yielding high rate of false positives and lack of specificity of measurement. Generalizing to all 1st Pg is inappropriate - restricting sample to only those women who had completed the Rotter I-E scale in 1979, effectively eliminated most (339 of 425) of the teenagers who had delivered; women in the pre-1980 DEL grp that was eliminated had the highest % exceeding CESD cut-off (34%) compared to pre-1980 AB (27%) and post-1980 AB (24%) & DEL (24%) grps. Limited to women under 38 years of age in 1992. Variable used to define race included nonBlack and nonHispanic minorities in the White category.

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Notes and Additional Limitations Specific to Study Listed:
Cogle, J.R., Reardon, D.C., & Coleman, P.K. (2002). Depression associated with abortion and childbirth: A long-term analysis of the NLSY cohort. <i>Medical Science Monitor</i> , 9, CR105-112.	Drawn from a larger subsample of 1,884 women with first abortion or first delivery between 1980 and 1992 and who had completed both the 1979 Rotter I-E scale and the 1992 CES-D scale; total AB & DEL grps not reported; average age figure based on 884 women (AB = 293; DEL = 591); subsample ns varied from 1031 - 1361 depending on the analyses.	AB N= 131 - 164 DEL N= 877 - 1197	% women exceeding the 1992 CESD cut-off score (>15).	<p>AB grp had higher % scoring ≥ 16 on CES-D in 1992 (27% vs. 21%), controlling for age, race, education, income, marital status, history of divorce, and abbreviated I-E score.</p> <p>AB group had higher depression risk among women who were White, married, and who did not have a 1st marriage ending in divorce, controlling for relevant covariates.</p> <p>Significant differences not found among Black/Hispanic women, unmarried women, or women with a 1st marriage ending in divorce, controlling for relevant covariates.</p>	This study is similarly designed and based on the women erroneously identified in first set of analyses in Reardon & Cogle (2002), except that women who had intended pregnancies are now added to DEL group, reducing % exceeding cut-off score. Reasons for discrepancies in AB & DEL groups from previous study not clear, possibly due to different covariates (age vs. age at 1st Pg) used in the two studies for unknown reasons. Average age based on 884 women so difficult to understand where ns exceeding that n in the regression analyses came from given age is a covariate in those analyses. Variable used to define race included non-Black and non-Hispanic minorities in the White category.

methodological limitations, including lack of information about pregnancy wantedness and lack of assessment of other critical variables known to co-vary with both pregnancy outcome and mental health (e.g., prior reproductive history, prior mental health problems, violence exposure, etc).

The largest and most methodologically rigorous Finland study used definitions provided by the American College of Gynecology (ACOG) to analyze direct pregnancy-related deaths (deaths occurring within one year of end of pregnancy from causes related to or aggravated by the pregnancy or its management,

but not from accidental or incidental causes) separately from pregnancy-associated (deaths occurring within one year from end of pregnancy, regardless of cause of death) (Gissler, Berg, Bouvier-Colle, & Buekens, 2004b). These analyses revealed that women in the abortion group had lower rates of pregnancy-related deaths than women in the delivery group (1.3 vs. 3.9 per 100,000 pregnancies), but higher rates of pregnancy-associated deaths. However, when therapeutic abortions were excluded from the category of pregnancy-associated deaths, women in the abortion group no longer had higher pregnancy-associated death rates than women in the delivery group.

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Notes and Additional Limitations Specific to Study Listed:
Schmiege, S., & Russo, N. F. (2005). Depression and unwanted first pregnancy: Longitudinal cohort study. <i>British Medical Journal</i> , 331, 1303-1305.	Two samples were drawn due to coding issues in the initial study; both the initial and corrected sample ns are reported here. 1. Initial sample: 1247 women identified as reporting a first unwanted PG between 1970 and 1992 that resulted in abortion (N=479) or delivery (N=768). 2. Corrected sample: 1744 women identified as reporting a first unwanted PG 1970 & 1992 that resulted in abortion (N=461) or delivery (N=1283). Results were similar in both samples & only results from corrected sample presented here.	1. AB N=479 DEL N=768 2. AB N=461 DEL N=1283	Both % women exceeding the 1992 CESD cutoff score (>15) and continuous 1992 CESD scores reported. Education, income, and family size also examined.	% exceeding cutoff score on 1992 CESD did not significantly differ for AB vs DEL groups, controlling for age at 1st Pg, race, marital status, education, and family income, in either the full sample (25% vs. 28%) or the post-1979 subsample (23% vs. 23%) for all women. AB sig. associated with lower education and income and larger family size, all risk factors for depression. Additional analyses published in response to debates over points of design did not change the pattern of results. The only sig. difference between AB & DEL grps found was in unadjusted analyses when subsequent abortions excluded from both groups (AB = 21% >15 vs. DEL = 28% >15); the difference was not sig. when covariates controlled. Findings did not vary across groups known to vary in underreporting, including married white women, unmarried White women, unmarried Black women, non-Catholics, and Catholics.	Note: NLSY staff provided coding to ensure proper identification of sample, but last line of code inadvertently omitted in initial analyses. Differs from other studies in focusing on unwanted 1st Pg. Study criticized for not controlling same variables as previous studies, resulting in a series of analyses, including those limited to post-1980 AB & DEL grps. Although underreporting bias a concern, findings did not differ among grps known to vary in such bias. Limited to women under 38 years of age in 1992.

This study affirms the importance of making a distinction between pregnancy-related and pregnancy-associated deaths in drawing valid conclusions about the association between abortion (vs. delivery)

and subsequent risk for various causes of death and also establishes the importance of separating therapeutic from elective abortions when attempting to draw such conclusions.

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Notes and Additional Limitations Specific to Study Listed:
Reardon, D., Coleman, P. K., & Cogle, J. R. (2004). Substance use associated with unintended pregnancy outcomes in the National Longitudinal Study of Youth. <i>American Journal of Drug and Alcohol Abuse</i> , 30, 369-383.	After excluding all women Pg before 1980, identified 1748 women reporting a first unintended PG between 1980 and 1988 that resulted in abortion (N=213) or delivery (N=535), or had never been Pg. (N= 1144); a subsample of women responded to alcohol questions; alcohol analyses appear to be based on 1243 women.	AB N= 213 DEL N= 535 Never Pg N= 1144	11 yes/no items related to alcohol abuse symptoms; 4 related to substance use (# days drank in last mo; # drinks consumed on days when drank; if ever used marijuana or cocaine in last mo).	Controlling for age, race, marital status, income, education, pre-Pg RSE and pre-Pg I-E, no sig. differences among groups on # of drinks; in % scoring 2 or more or % scoring 4 or more on items related to alcohol abuse; in the number of drinks consumed, or in the use of cocaine. AB grp drank sig. more days in last mo (6.36) than DEL grp (4.79) but not than Nev Pg grp (5.93); and were more likely to use marijuana in last month (18.6%) than the DEL or Nev Pg grps (7.9%).	Exclusion of women Pg before 1980 makes sample unrepresentative and generalization to unintended first Pg inappropriate as noted above. The large number of tests performed, single item measures of key dependent variables, and small magnitude of effects limit conclusions that can be drawn from this study. Drinking on an average of 6.36 (AB) vs. 4.79 (DEL) days per mo. not indicator of clinically significant alcohol abuse. Variable used to define race included nonBlack and nonHispanic minorities in the White category.

The most consistent findings across the Medi-Cal and Finland record-based studies were the higher rates of violent death for women in the abortion group. In the Finland study described above, women in the abortion group had higher rates of violent pregnancy-associated deaths, and a higher proportion of their overall pregnancy-associated deaths were due to violent causes (Gissler et al., 2004b). In interpreting this finding, it is useful to recall the distinction between risk and cause discussed above. Abortion is a marker of risk for violence, not a cause of violence. Thus it is important to control for violence exposure in studies of pregnancy outcome.

Secondary analyses of survey data. Fifteen papers based on secondary analyses met inclusion criteria for our review. These were based on nine data sets. Eight data sets were from the United States: Five were based on U.S. national probability surveys, and three were based on local metropolitan area surveys. One paper

was based on analyses of the longitudinal New Zealand Christchurch Health and Development survey. Key findings and methodological limitations of these studies are summarized in Table 2.

National Longitudinal Survey of Youth (NLSY). The NLSY has been the data set used most frequently to examine the relationship of abortion to mental health outcomes. The NLSY is a longitudinal national survey of a cohort of males and females aged 14-21 years in 1979. Papers meeting our inclusion criteria assessed the following outcome variables: self-esteem measured in 1987 (2 studies), risk for depression measured in 1992 (3 studies), and substance use measured in 1988 (1 study). This set of papers demonstrates the problems of trying to base conclusions about the mental health effects of abortion on secondary analyses of data sets collected for other purposes. Conclusions of researchers

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

National Longitudinal Study of Adolescent Health (ADD-HEALTH)

General Description: ADD-HEALTH uses a multi-stage, school-based, longitudinal design in which data were three times: initial (1994-1995), and approximately 1 year (1996), and 6 years (2001-2002) later. At Wave I all participants were in grades 7-12. All Wave I (N= 90,118) completed an in-school questionnaire; a subsample (N=12,105) completed an additional computer-assisted in-home interview that included questions about sexual history and religion. This subsample was chosen by identifying a group of students who were representative of the adolescent population in grades 7-12 during the 1994-1995 school year; in addition, adolescents who were disabled, African American students from well-educated families, Chinese, Cuban, Puerto Rican, living with twin, living with a full sibling, living with a half sibling, living with a non-related adolescent, and siblings of twins were oversampled.

Limitations Common to All Studies Based on this Data Set: School-based population does not include students who drop out due to Pg; ethnic minorities in sample may be particularly unrepresentative of the adolescent population as a whole. 1-item measures psychometrically weak.

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Key Findings	Notes and Additional Limita- tions Specific to Study Listed:
Coleman, P.K. (2006). Resolution of unwanted pregnancy during adolescence through abortion versus childbirth: Individual and family predictors and consequences. <i>Journal of Youth and Adolescence</i> , 35, 903-911.	130 adolescents in grades 7-11 who completed both Waves I & II and experienced a Pg, described as "not wanted" or "probably not wanted."	AB N = 65 DEL N = 65	Single-item measures of counseling, 12-month trouble sleeping, 30-day cigarette use, 30-day marijuana use, 12-month alcohol use, problems with parents and with school due to alcohol use.	Controlling for risk taking and desire to leave home, AB group more likely to have counseling, trouble sleeping, and use marijuana in past 30 days (problems with parents due to alcohol use approached significance).	Number of total pregnancies unknown, but small n's raise questions about underreporting and drop-out rates. Single item outcome measures psychometrically weak. Percentages and ns for outcome variables not reported so frequency of problem unknown; previous mental health problems not controlled. Given the large number of variables in the data set, why these particular variables were included is unclear.

analyzing this same data set and even the same dependent variable varied markedly depending on sampling and analytic strategy.

Self-esteem. The first of the abortion studies to be based on this data set focused on self-esteem as measured by the Rosenberg self-esteem scale (RSE; Rosenberg, 1965). This first study (Russo & Zierk, 1992) analyzed a total sample of 5,295 women (773 of whom reported having at least one abortion). Women who had an abortion had mean RSE scores comparable to those of all women (33.3 vs. 33.2, respectively); women who had one abortion also had significantly higher RSE in 1987 than the other

two groups (women with no abortions, women with repeat abortions), although the relationship was extremely small. When contextual variables were controlled (education, income, employment, marriage, number of children, whether the pregnancy was wanted or unwanted), however, neither having one abortion nor repeat abortions was related to subsequent self-esteem. After eliminating from the study women who had an abortion before RSE was measured in 1980, further analyses found that preexisting self-esteem was the most important predictor of 1987 RSE, followed by having more education, higher income, employment, and fewer children.

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

National Survey of Family Growth (NSFG)

General Description: The NSFG Cycle V sample is a subsample of 10,847 women aged 15-44 drawn from the larger national probability sample of the National Health Interview Survey. The NSFG is thus a stratified, multistage design involving individual sampling rates that requires using sampling weights in computing statistics.

Limitations Common to All Studies Based on this Data Set: Retrospective data that may involve recall of events occurring decades previously.

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Key Findings	Additional Limitations Specific to Study Listed:
Cogle, J., Reardon, D.C., Coleman, P.K., & Rue, V.M. (2005). Generalized anxiety associated with unintended pregnancy: A cohort study of the 1995 National Survey of Family Growth. <i>Journal of Anxiety Disorders</i> , 19, 137-142	Study sample: (1) all women having an unintended Pg ending in abortion for their first Pg event and (2) all women having an unintended Pg ending in live birth delivery for their first Pg event who had no abortions after that Pg. Women who experienced a prolonged period of anxiety previous to or at the same age as the Pg event were excluded from the sample.	AB N = 1033 DEL N = 1813	Dichotomous measure (yes/no) of generalized anxiety (GE)	Significantly higher rate of GE in abortion vs. delivery group (13.7% vs. 10.1%), controlling for race and age at interview. In stratified sub-analyses, difference sig. for unmarried or under 20 at 1st Pg, but not for married women.	Women reporting pre-Pg anxiety excluded so cannot generalize to all first unintended pregnancies; misleading language implies generalized anxiety disorder (GAD) is assessed, but items used to construct generalized anxiety variable are not congruent with DSM definitions of generalized anxiety disorder, making clinical implications problematic; differential exclusion from women with subsequent abortions from delivery but not abortion group; sampling weights not used in statistical analyses; stratification used rather than controlling for relevant variables; analyses not conducted to determine the contribution of abortion to variance over and above other relevant predictor variables.

This study reported a number of relationships that have implications for what should be controlled when analyzing NLSY data, especially the importance of controlling for wantedness of pregnancy and separating women with one abortion from those having repeat abortions. The number of abortions was slightly but significantly and positively correlated with unwanted births

($r = .11$). Furthermore, repeated unwanted pregnancy, regardless of pregnancy outcome (birth or abortion), was significantly correlated with greater likelihood of living in poverty ($r = .15$) and lower education ($-.13$).¹

Depression risk. Using a very different approach, three studies focused on the effects of first pregnancy

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Commonwealth Fund Health of American Women Survey

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Key Findings	Additional Limita- tions Specific to Study Listed:
Russo N., & Denious, J. (2001). Violence in the lives of women having abortions: Implications for practice and public policy. <i>Professional Psychology: Research and Practice</i> , 32, 142-150.	Secondary analyses a random household telephone survey of over 2,500 women and 1,000 men aged 18 or over and residing in the continental U.S., conducted in 1993. Analyses based on responses of 2,525 women, 324 of them identified as having had at least 1 abortion; ns varied depending on missing data.	AB (N= 324) Others (N= 2,201)	Global self esteem; abbreviated CES-D (6 items); 1-item measures of suicidal ideation in past year; if told by doctor she had anxiety/depression in past 5 years, 1-item life satisfaction measure	AB correlated positively with CESD (.08), having suicidal thoughts (.08), being told by a doctor had anxiety/depression (.08) & negatively with life satisfaction (-.06). Also correlated with experiencing rape (.06), childhood physical (.15) & sexual (.18) abuse, having a violent partner (.11), & a partner who refused to use condom (.06). Controlling for race, education, children living at home, marital status, and partner and violence variables, abortion not significantly related to any outcome variable.	Outcome and violence measures psychometrically weak. Timing of events vis-a-vis abortion unknown. Only women married or living as a couple were asked about partner violence. Limited generalizability of study group: have telephone, younger teenagers not included, older age (median 40-44), 57% married. Low reported abortion rate (13%) could reflect underreporting and/or recall bias. Only one question asked about abortion history; repeat abortions not identified. Comparison is with other women, not women with unintended Pg.

outcome (abortion vs. delivery) on risk for subsequent depression (measured in 1992 by the Center for Epidemiological Studies-Depression scale (CES-D; Radloff, 1977). Reardon and Cogle (2002a) focused on unintended first pregnancy outcome (abortion vs. delivery). After correcting an initial coding error, they reported analyses controlling for age at first pregnancy, race, marital status, and whether the woman was in her first marriage. They also attempted to control for prior mental health by including only women who had completed an abbreviated Internal-External Locus of Control scale (I-E Scale; Rotter, 1966), assessed in 1979, prior to having a first pregnancy. Among all women, 25% of the delivery group exceeded the CES-D cutoff score for depression (>15) compared to 27% of the abortion group, a nonsignificant difference. Among married women in this subsample, a significantly higher percentage of women in the abortion group (26%) than in the deliv-

ery group (19%) exceeded the CES-D cutoff score. Among unmarried women in this subsample, the findings were reversed, although not statistically significant (36% vs. 29%).

Cogle et al. (2003) published another paper also focusing on first- pregnancy outcome (abortion vs. delivery) relative to the same outcome variable, 1992 CES-D. This study is based on essentially the same sample as the previous one with the primary difference being that women with wanted pregnancies were also included in the delivery group. Again, a larger percentage of women in the abortion group exceeded the CES-D cutoff score for depression compared with women in the delivery group.

Both of these studies are characterized by a number of problems, the most important of which are the

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Washington, DC, Metropolitan Area Drug Study

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Key Findings	Additional Limita- tions Specific to Study Listed:
Coleman, P., Reardon, D.C., & Coughle, J. R. (2005). Substance use among pregnant women in the context of previous reproductive loss and desire for current pregnancy. <i>British Journal of Health Psychology</i> , 10, 255-268	Data drawn from the public release data set that resulted from the Washington, D.C. Metropolitan Area Drug Study (CD*MA DS). The initial sample, constructed to oversample for low birth weight, pre-term, and admitted maternal drug use, consisted of 1,020 women giving birth in Washington, DC area hospitals in 1992. The initial sample was predominantly never married, Black, between 19 and 34 years of age, high school or less education, and of relatively low family income (under \$20,000). Of these cases, those with known medical outcomes of previous pregnancies were selected for further analysis.	Sample sizes for the several reported analyses differ from one analysis to another. The key comparisons reported in Table 3, in which the odds ratios for drug use during the current pregnancy as a function of prior abortion history seems to be based upon comparisons of 144 women who reported no prior abortions and 282 women who reported one or more abortions prior to the index delivery. (These numbers were not directly reported in the paper but were determined through an examination of the public release data set used in these analyses. The numbers are essentially consistent with percentages and methods reported in the paper.)	Differential odds ratios for the use of marijuana, cigarettes, alcohol, crack cocaine, other cocaine, and any illicit drugs are reported for 1 previous abortion vs no abortion history and 2 or more abortions vs no abortion history after statistical adjustment for number of prior births, miscarriages, and still births; age; education; number of people the respondent lives with; and a binary indicator reflecting if prenatal care was sought in the first trimester.	Adjusted for covariates, a statistically higher odds ratio was reported for the use of legal and illegal substances during the index pregnancy if the woman had a prior history of abortion.	The sample very specialized. No indication that sampling fractions used in analysis to reweight sample. Many of the illegal substance categories are fairly rare (e.g., there are only 58 cases of any reported crack cocaine use during Pg among the subset of cases who had usable data on abortion history.) Results look very different for covariate adjusted analyses and unadjusted analyses. No regression diagnostic results are reported.
Coleman, P., Maxey, C.D., Rue, V.M., & Coyle, C.T. (2005). Associations between voluntary and involuntary forms of perinatal loss and child maltreatment among low income mothers. <i>Acta Paediatrica</i> , 94, 1476-1483.	Data drawn from Fertility and Contraception Among Low Income Child Abusing and Neglecting Mothers in Baltimore, MD, 1984-1985, a study of family patterns and contraceptive use among maltreating mothers. Sample of 518 mothers (Age range 18-50; 79% Black; 6.8% employed) who were receiving AFDC. All women interviewed in home	118 physically abusive mothers and 119 neglecting mothers selected from cohort receiving Child Protective Services (CPS) and 281 mothers without maltreatment offences. In interview, 100 women reported 1 abortion, 59 reported 2+ (abortion ave 6.5 years earlier), 99 reported 1 miscarriage or stillbirth, 34 reported 2+ (ave 7.1 yrs earlier).	Association between self-reported abortion or miscarriage/stillbirth history and being in the physically abusing or neglecting groups. Logistic analyses controlled for covariates (single-item measures) associated with maltreatment (e.g., more children, history of depression, worries about income, etc).	Adjusted for covariates, women reporting 1 abortion were not more likely than those reporting no abortions to be in child neglect group, but were sig more likely to be in physical abuse group. History of multiple induced abortions not related to increased risk for either abuse or neglect. Maternal history of multiple miscarriages and/or stillbirths compared to no history was associated with increased risk of physical abuse and neglect.	Retrospective self-reports of abortion in interview unreliable. Abortion likely underreported. Sample not representative of U.S. women. No info about nature of abortion. Single-item measures of covariates. Causal direction ambiguous. Same factors (e.g., poverty; drug use) may contribute to increased risk of child maltreatment and abortion.

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Washington, D.C. Metropolitan Area Drug Study

Citation	Data Source/ Population Studied	Sample Sizes	Controls/ Covariates	Primary Outcome	Results
Coleman PK, Reardon D.C., Rue, V.M., & Cougle, J. (2002). A history of induced abortion in relation to substance use during subsequent pregnancies carried to term. <i>American Journal of Obstetrics and Gynecology</i> , 187, 1673-1678.	Data drawn from the National Pregnancy and Health Survey conducted in 1992 whose purpose was to assess drug and alcohol consumption in a national sample of pregnant women (N= 2,613). Hospitals with < 200 annual births were selected in the first stage of sample selection; individual mothers within hospitals were randomly selected in the second stage. Soon after delivery women were interviewed about reproductive history and completed a drug use questionnaire answer sheet in response to interviewer questions. Samples used in analyses were limited to women who recently had given birth, and had one previous induced abortion, one previous birth or no previous births or abortions.	The primary sample of women with a recent delivery (N= 607) has two subgroups: 74 women with one previous induced abortion and 531 women with one previous birth. The secondary sample included 738 first-time mothers with no previous abortions. Both groups were primarily White, married, and employed full-time. The average age of the two groups respectively was 26.5 and 23.4 years.	Association between previous reproductive outcome and usage of alcohol or illicit drugs during most recent pregnancy. Differential odds rates for use of any illicit drugs, marijuana, cigarettes and alcohol reported for 1 previous abortion vs. 1 previous birth group, and 1 previous abortion vs. first birth group. Adjusted for covariates by stratifying covariates related to substance use type and running separate analyses.	Women with a previous abortion had higher rates of any illicit drug use, marijuana use and alcohol use than women with a previous live birth. Differences between reproductive history groups appeared greater when time since previous pregnancy was longer (3-5 vs. < 2 years). The abortion group also reported higher rates of illicit drug use, marijuana, and alcohol use than first-time mothers.	Samples analyzed not representative of total NPHS sample or of U.S. women giving birth. Retrospective self-reports of abortion may be unreliable. Abortion likely underreported. Single-item outcome measures. No statistical adjustment for number of significance tests. Confounds not controlled. Small size of abortion group led to many cell counts <5 in subgroup analyses which were intended to control for confounds. Differences found could be due to other unmeasured factors such as whether pregnancy intended, domestic violence or sexual abuse. Comparisons between previous abortion and previous birth groups could be explained by child-care demands on mothers or differential stress of first versus later completed pregnancy.

miscoding of the first pregnancy variable and the differential exclusion of women having subsequent abortions only from the delivery group (see Table 2 for details).

In an effort to redress these problems, Schmiede and Russo (2005) reexamined depression risk in the NLSY. Using codes provided by the NLSY staff, they identified a sample of 1744 women as having an unwanted first pregnancy. (They, too, had a coding error in their initial article, but it did not affect the pattern and sig-

nificance of their findings when corrected. After a series of interchanges in which they addressed criticisms of their approach, we report here the findings based on the corrected codes verified by the NLSY staff and published with the analyses.) First, Schmiede and Russo found that the sampling strategy that Reardon and Cougle (2002a) and Cougle et al. (2003) had used to control for prepregnancy psychological state (which was to include only those women who had completed the Rotter I-E scale in 1979 prior to their first pregnancy) resulted in excluding from their sample the

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
INTERNATIONAL STUDIES–NEW ZEALAND

Christchurch Health and Development Study

General Description: The Christchurch Health and Development Study a longitudinal study of a cohort of 1,265 children born in 1977 in the Christchurch, New Zealand, urban region who were studied from birth to age 25, including 630 females. Information was obtained on: (a) the history of PG/abortion for female participants over the interval from 15-25 years; (b) measures of DSM-IV mental disorders (including major depression, overanxious disorder, generalised anxiety disorder, social phobia, & simple phobia), and suicidal behaviour for intervals 15-18, 18-21 and 21-25 years; and (c) childhood, family and related confounding factors.

Limitations Common to All Studies Based on this Data Set: Common to All Studies Based on this Data Set: Neither intendedness nor wantedness of Pg controlled; in New Zealand to obtain a legal abortion, a woman is referred to two specialist consultants by her doctor; the consultants must agree that either (1) the Pg would seriously harm the life or the physical or mental health of the woman or baby; (2) the Pg is the result of incest; or (3) the woman is severely mentally handicapped. An abortion will also be considered on the basis of age or when the Pg is the result of rape. Comparisons with population data suggest abortion is underreported. Measures of child abuse psychometrically weak and it is likely underreported.

New Zealand	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Key Findings	Additional Limitations Specific to Study Listed:
Fergusson D.M., Horwood, L.J., & Ridder, E.M. (2006). Abortion in young women and subsequent mental health. <i>Journal of Child Psychology & Psychiatry</i> , 47, 16-24.	Forty-one percent of women Pg on at least one occasion prior to age 25; 14.6% have at least one abortion. Sample sizes in analyses ranged from 506 and 520 depending on the timing of assessment. Ns for prospective analyses were provided in personal communication from the author.	Concurrent analyses: AB N= 74 DEL N= 131 Never Pg N= 301 Prospective analysis: AB N= 48 DEL N= 77 Never Pg N= 367	In concurrent analyses, yes/no diagnosis of major depression, anxiety disorder, alcohol and illicit drug dependence, suicidal ideation in previous 12 mo., and total # of disorders. In prospective analysis, total number of disorders from 21-25 yrs.	In concurrent analyses, controlling for covariates, AB grp had sig (p<0.05) higher rates of depression, suicidal ideation, illicit drug dependence, & total mental health problems than the DEL grp & except for alcohol and anxiety disorder, significantly higher rates of disorder than the Never Pg grp. A prospective analysis used Pg/abortion history prior to age 21 to predict mental health outcomes from 21-25 years. Similarly, after covariate adjustment, the AB grp had a sig. higher total # of disorders than the other grps, which did not sig differ from each other.	Although a longitudinal study, most results reported involved the concurrent assessment of Pg status and mental health. The one prospective analysis was limited to number of disorders owing to the relatively sparse data for specific disorders over the interval 21-25 years and the smaller number of women who became pregnant by age 21.

women who had the highest risk for depression—those who had delivered at a younger age. Significantly more women who had delivered pre-1980 exceeded the CESD cutoff score (33.5%) than who had an abortion pre-1980 (26.5%). Like Cougle et al. (2003), they controlled for age of first pregnancy, race, education, and family income. However, instead of excluding women based on previous marriage, they considered it more

appropriate to maximize generalizability by controlling for marital status. When Schmiede and Russo analyzed the full sample (not restricted on the basis of I-E scores), they found no significant differences in depression between the abortion and delivery groups when race, age at first pregnancy, 1992 marital status, education, and family income were controlled: 28.3% of women in the delivery group exceeded the CESD

Table 1B: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
INTERNATIONAL STUDIES–NEW ZEALAND (continued)

New Zealand	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Key Findings	Additional Limitations Specific to Study Listed:
Fergusson, D.M., Boden, J.M., & Harwood, L.J. (2007). Abortion among young women and subsequent life outcomes. <i>Perspectives on Sexual and Reproductive Health</i> , 39, 6-12.	492 women for whom full information on Pg history, education, income, welfare dependence, employment and partnership variables to age 25 was available classified in 3 groups: abortion before age 21 (AB); Pg but no abortion age 21 (DEL)(77); and never Pg before age 21 (Never Pg); 125 had had at least one Pg by age 21; of 172 Pg reported, 55% ended with live birth, 31% by abortion, & 14% in miscarriage. Two sets of analyses: (1) one based on 1st Pg outcomes, AB vs DEL; (2) Pg-no abortion vs. Pg with abortion as correlated dichotomous predictor variables to take into account possible overlap between abortion and Pg without abortion.	AB N= 48 DEL N= 77 Never Pg N= 367	Social and economic outcomes at ages 21-25; 4 educational variables; family income, welfare dependence, employment, partner violence (items from the Conflict Tactics Scale), relationship quality (items from Intimate Relations Scale) & relationship satisfaction.	AB grp sig more likely than DEL grp to have attended university, gained a university degree, & gained a tertiary qualification other than a university degree, & less likely to have been welfare-dependent. Also had sig higher mean personal income & experienced sig. lower mean level of partner violence. AB grp not sig different from Never Pg group on all education outcomes, mean family income, and both partnership measures. Women in the DEL grp had sig lower intelligence scores and levels of educational achievement in childhood & were more likely to drop out of school. Most differences explained by pre-Pg family, social and educational characteristics, except AB grp continued to have sig higher levels of subsequent educational achievement than DEL grp. For all outcomes, DEL grp fared sig less well than Never Pg grp. The pattern of results was similar across the two forms of analysis.	Comparisons based on relatively small numbers of women.

Notes: AB = Abortion group, DEL = delivery group, Pg = pregnancy

cutoff score compared to 25% of the abortion group, a nonsignificant difference.

They also examined the implications of the practice of differentially excluding all women who had subse-

quent abortions from only the delivery group (but not from the abortion group) by comparing abortion and delivery groups with women having subsequent abortions excluded from both groups. Using this approach, significantly more women in the delivery group

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES

National Longitudinal Survey of Youth (NLSY)

General Description: The National Longitudinal Survey of Youth (NLSY) is based on annual interviews with a stratified, multi-stage national probability sample of noninstitutionalized civilian men and women aged 14-21 as of 1979, with oversampling of Blacks, Hispanics, and poor Whites. Relevant measures include: an abbreviated version of the Rotter internal-external locus of control scale (IE, Rotter, 1966; assessed in 1979); global self esteem (RSE, Rosenberg, 1979; assessed in 1980 & 1987); Center for Epidemiological Studies-Depression Scale (CES-D, Radloff, 1977; assessed in 1992); reproductive histories were first taken in 1982 and updated every 2 years subsequently.

Limitations Common to All Studies Based on this Data Set: No study used sampling weights so that normative statements are inappropriate and alpha levels are likely to be elevated, increasing probability of identifying difference due to chance as a reliable difference. Underreporting of abortion raises question of possible reporting bias, but direction of reporting bias unclear as women may be less likely to report stigmatized experiences (having an abortion, mental problems, experiencing violence), but those who are willing to report one stigmatized condition may be more willing to report others, increasing the likelihood of finding a correlation between 2 stigmatized events. Ns of analyses vary depending on covariates and are not always clear. Large sample sizes mean that small effects are statistically significant. CES-D controversial due to cutoff at > 15 yielding high rate of false positives and lack of specificity of measurement. Generalization limited to restricted age range (women 14-24 in 1979).

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Results	Notes and Additional Limita- tions Specific to Study Listed:
Russo, N.F., & Zierk, K.L. (1992). Abortion, child-bearing, and women's well-being. <i>Professional Psychology: Research and Practice</i> , 23, 269-280.	1. 5,295 women for whom there were NLSY interviews involving the assessment of well-being in 1987; 773 had at least one abortion; 233 had repeat abortions. 2. Additional analyses based on 4502 women who had no abortions before their 1980 interview.	1. AB N = 733 Other N = 4562 2. AB N = 317 Other N = 4185	1987 Global self-esteem (RSE)	MRSE = 33.2 & 33.3 for all women vs. women having at least 1 abortion; 1. Women who had 1 abortion had higher RSE than no abortion or multiple abortion groups; when childbearing and resource variables were controlled, neither having 1 abortion nor having repeat abortions were significantly related to RSE. Total abortions correlated with total unwanted births ($r=.11$). 2. in subsample 1980 RSE was the strongest predictor of 1987 SE (partial $r=.38$).	No clinical cut-off score for RSE & clinical significance of scores is unknown; large sample means small effects statistically significant. Age range of sample limited to women 22-33 in 1987.
Russo, N.F., & Dabul, A.J. (1997) The relationship of abortion to well-being: Do race and religion make a difference? <i>Professional Psychology: Research and Practice</i> , 28, 23-31.	1. 4913 women drawn from the sample of 5,295 women described above (3572 White & 1341 Black); 721 had at least one abortion, 175 had repeat abortions. 2. Additional analyses based on 4336 women (3,147 White & 1,189 Black) who had no abortions prior to 1980 interview.	1. AB N = 721 Other N = 4192 2. AB N = 317 Other N = 4502	1987 Global self-esteem (RSE)	Primary findings did not vary across groups known to vary in underreporting. 1. When childbearing and resource variables were controlled, neither having 1 abortion nor having repeat abortions significantly related to RSE, regardless of race or religion. 2. 1980 SE was the strongest predictor of 1987 SE (partial $r=.39-.42$) regardless of race or religion.	Religion measured in 1979 only; highly committed fundamentalist women not identified; sample does not include Asians or Native Americans. Age range of sample limited to women 22-33 in 1987.

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Notes and Additional Limitations Specific to Study Listed:
Reardon, D.C., & Cougle, J.R. (2002). Depression and unintended pregnancy in the National Longitudinal Survey of Youth: A cohort study. <i>British Medical Journal</i> , 324, 151-152.	Two samples were drawn due to coding issues in the initial study; both the initial and corrected sample ns are reported here. 1. Initial sample: 421 women identified as reporting a first unintended Pg between 1980 and 1992 that resulted in abortion (N=293) or delivery with no subsequent history of abortion in the delivery grp (N=128). 2. Corrected sample: 1076 women identified as reporting a first unintended Pg between 1980 and 1992 that resulted in abortion (N=293) or delivery with no subsequent history of abortion in the delivery grp (N=783). Results were similar in both samples & only results from corrected sample are presented here.	1. AB N=293 DEL N=128 2. AB N=293 DEL N=783	Percent of women exceeding the 1992 CES-D cut-off score (>15).	AB grp had higher % scoring >15 on CES-D in 1992 (27% vs. 25%), controlling for family income, education, race, age at 1st Pg, and 1979 I-E score. Sig higher risk for AB grp among married women (26% vs. 19%), but not among unmarried women (29% vs. 36%), controlling for family income, education, race, age at 1st Pg, and 1979 I-E score.	Note: Differs from RSE studies in focusing on outcome of 1st Pg. Subsequent reanalysis by Schmiege & Russo (2005) showed that findings in corrected sample still based on miscoded data. Excluded women with subsequent history of abortion only from the delivery grp. Used I-E score as a control for pre-existing mental health but scale is not a measure of mental health. Generalizing to all 1st Pg is inappropriate – sample restricted to only women who had completed the Rotter I-E scale in 1979, effectively eliminating most (339 of 425) of the teenagers who had delivered; women in the pre-1980 DEL grp that was eliminated had the highest % exceeding CES-D cut-off (34%) compared to pre-1980 AB (27%) and post-1980 AB (24%) & DEL (24%) grps. Variable used to define race included nonBlack and nonHispanic minorities in the White category. Age range of sample limited to women 27-38 in 1992.

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Results	Notes and Additional Limita- tions Specific to Study Listed:
Cogle, J.R., Reardon, D.C., & Coleman, P.K. (2003). Depression associated with abortion and childbirth: A long-term analysis of the NLSY cohort. <i>Medical Science Monitor</i> , 9, CR105-112.	Based on a larger subsample of 1,884 women with first abortion or first delivery with no subsequent abortions between 1980 and 1992 and who had completed both the 1979 Rotter I-E scale and the 1992 CES-D scale; total AB & DEL grps not reported, but reports an average age figure based on 884 women (AB = 293; DEL = 591); subsample ns reported as varying from 1031-1361 depending on the analyses.	AB N= 131 - 164 DEL N= 877 - 1197	Percent of women exceeding the 1992 CES-D cutoff score (>15).	Final corrected table: AB grp had higher % scoring >15 on CES-D in 1992, controlling for age, race, education, income, and abbreviated I-E score. Higher depression risk found for AB group among women who were White, married, and who did not have a first marriage ending in divorce, controlling for relevant covariates. Sig differences not found among Black/Hispanic women, unmarried women, or women with a first marriage ending in divorce, controlling for relevant covariates.	This study is similarly designed and based on the women erroneously identified in first set of analyses in Reardon & Cogle (2002), except that women who had intended pregnancies are now added to DEL group, reducing % exceeding cut-off score. Reasons for discrepancies in AB & DEL groups from previous study not clear, possibly due to different covariates (age vs. age at 1st Pg) used in the two studies for unknown reasons. Average age figure based on 884 women so not clear how ns in the regression analyses determined, given they exceed that number and age is a covariate in those analyses. Variable used to define race included non-Black and non-Hispanic minorities in the White category. Age range of sample limited to women 27-38 in 1992.

(28.1%) than the abortion group (20.7%) exceeded the CESD cutoff score ($p < .01$). These analyses illustrate that the sampling and exclusion strategies researchers use to analyze secondary data sets can dramatically alter the conclusions reached regarding the relative risks for depression accompanying childbirth versus abortion. When attempting to examine the effects of first pregnancy outcome, it is important to control for both number of subsequent abortions and number of subsequent births in both groups.

Substance use. Reardon et al. (2004) used NLSY data to examine substance abuse among 535 women who had terminated a first unintended pregnancy compared with 213 women who had delivered a first unintended pregnancy and 1144 women who had never been pregnant. These researchers again excluded women pregnant before 1980 (i.e., those known to be at a significantly higher risk for depression than other women in the sample and more likely to be found in the delivery group; Schmiede & Russo, 2005). They

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Notes and Additional Limitations Specific to Study Listed:
Schmiege, S., & Russo, N. F. (2005). Depression and unwanted first pregnancy: Longitudinal cohort study. <i>British Medical Journal</i> , 331, 1303-1305.	Two samples were drawn due to coding issues. In the initial study, both the initial and corrected sample ns are reported here. 1. Initial sample: 1247 women identified as reporting a first unwanted Pg between 1970 and 1992 that resulted in abortion (N=479) or delivery (N=768). 2. Corrected sample: 1744 women identified as reporting a first unwanted Pg between 1970 & 1992 that resulted in abortion (N=461) or delivery (N=1283). Results were similar in both samples and only results from corrected sample are presented here.	1. AB N=479 DEL N=768 2. AB N=461 DEL N=1283	Both % women exceeding the 1992 CES-D cutoff score (>15) and continuous 1992 CES-D scores reported. Education, income, and family size examined as outcomes.	Percentage exceeding cutoff score on 1992 CES-D did not sig differ for AB vs. DEL groups, controlling for age at 1st Pg, race, education, marital status, and family income, in full sample (25% vs. 28%) or post-1979 subsample (23% vs. 23%). AB sig associated with lower education and income and larger family size. Additional analyses published in response to debates over points of design did not change the pattern of results, with only sig difference found between AB & DEL grps in unadjusted analyses when subsequent abortions excluded from both groups; risk was lower in the AB grp (AB = 21% >15 vs. DEL = 28% >15); the difference was not sig when covariates controlled. Patterns of findings similar across groups known to vary in underreporting. Women who refused to fill out the confidential abortion card had sig lower CES-D scores than women who completed the card (13% vs. 25%).	Note: NLSY staff provided coding to ensure proper identification of sample, but last line of code inadvertently omitted in initial analyses, subsequently corrected. Differs from other studies in focusing on unwanted first Pg. Study criticized for not controlling same variables as previous studies, resulting in publication of a series of analyses, including those limited to post-1980 AB & DEL grps. Although underreporting bias a concern, the pattern of findings did not differ among grps known to vary in underreporting. However, lower CES-D scores among women who refused to fill out the confidential abortion card suggests that depression might be overestimated in the abortion group. Age range of sample limited to women aged 27-38 years in 1992.

also excluded women who had subsequent abortions from only the delivery group. In this subsample, controlling for prepregnancy I-E and RSE, age, race, marital status, income, and education, few significant

differences were found between groups in reported substance use. The exceptions were that women in the abortion group reported drinking on more days in the last month than the delivery group (6.4 vs. 4.8), but

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Results	Notes and Additional Limita- tions Specific to Study Listed:
Reardon, D., Coleman, P.K., & Cogle, J.R. (2004). Substance use associated with unintended pregnancy outcomes in the National Longitudinal Study of Youth. <i>American Journal of Drug and Alcohol Abuse</i> , 30, 369-383.	After excluding all women Pg before 1980, identified 1748 women reporting a first unintended Pg between 1980 and 1988 that resulted in abortion (N=213) or delivery (N=535), or had never been Pg (N=1144); a subsample of women responded to alcohol questions; alcohol analyses appear to be based on 1243 women.	AB N=213 DEL N=535 Never Pg N=1144	Eleven yes/no items related to alcohol abuse symptoms; 4 related to substance use (# days drank in last mo; # drinks consumed on days when drank; if ever used marijuana or cocaine in last mo).	Controlling for age, race, marital status, income, education, pre-Pg RSE and pre-Pg Rotter I-E score, no sig differences among groups on # of drinks; in % scoring 2 or more, or % scoring 4 or more on items related to alcohol abuse; in the number of drinks consumed, or in the use of cocaine. AB grp drank sig more days in last mo (6.36) than DEL grp (4.79) but not than Never Pg grp (5.93); and were more likely to use marijuana in last month (18.6%) than the DEL or Never Pg grps (7.9%).	Exclusion of women Pg before 1980 makes sample unrepresentative and generalization to unintended first Pg inappropriate as noted above. The large number of tests performed, single-item measures of key dependent variables, and small magnitude of effects limit conclusions that can be drawn from this study. Drinking on an average of 6.36 (AB) vs. 4.79 (DEL) days per mo. not indicator of clinically significant alcohol abuse. Variable used to define race included non-Black and non-Hispanic minorities in the White category.

not on more days than the never pregnant group (5.9%). They were also more likely to report using marijuana in the last month (18.6%) than did women in the delivery (7.9%) or never pregnant (7.9%) groups. These researchers did not control for history of drug use prior to the first pregnancy in their analyses despite the availability of this information in the data set and despite published findings in the literature that linked such drug abuse to later reproductive outcomes including likelihood of having an abortion (Mensch & Kandel, 1992; Rosenbaum & Kandel, 1990).

Evaluation of NLSY studies. Conclusions drawn from the NLSY about the mental health effects associated with abortion vary markedly by analytical strategy. Although the design of NLSY is longitudinal, like all survey data, it is correlational, making causal claims inappropriate. Collectively, these studies have a num-

ber of methodological limitations beyond those described above that make it difficult, if not impossible, to interpret the meaning of the correlations that are reported (see Table 2). Perhaps most importantly, none of these studies adequately controls for preexisting mental health or other important co-occurring risk factors prior to abortion or delivery (the Rotter I-E is not a measure of prior mental health), making it difficult to interpret the meaning of correlations observed between abortion and a mental health outcome. Covariates included in analyses varied across studies for unspecified reasons. Likewise, some contextual variables, such as marital status, that were shown in some studies to moderate results were not examined as moderators in other studies, compounding difficulties of comparing across studies. Further, some variables that were present in the NLSY and known to be related to the outcome variable under consideration (e.g., prior substance abuse) were omitted as covari-

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

National Longitudinal Study of Adolescent Health (ADD-HEALTH)

General Description: ADD-HEALTH uses a multi-stage, school-based, longitudinal design which collected data in three waves: initial (1994-1995), and approximately 1 year (1996), and 6 years (2001-2002) later. At Wave I all participants were in grades 7-12. All Wave I (N= 90,118) completed an in-school questionnaire; a subsample (N=12,105) completed an additional computer-assisted in-home interview that included questions about sexual history and religion. This subsample was chosen by identifying a group of students who were representative of the adolescent population in grades 7-12 during the 1994-1995 school year; in addition, adolescents who were disabled, African American students from well-educated families, Chinese, Cuban, Puerto Rican, living with twin, living with a full sibling, living with a half sibling, living with a nonrelated adolescent, and siblings of twins were oversampled.

Limitations Common to All Studies Based on this Data Set: School-based population does not include students who drop out due to Pg; ethnic minorities in sample may be particularly unrepresentative of ethnic minorities in the adolescent population as a whole. 1-item measures psychometrically weak.

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Notes and Additional Limitations Specific to Study Listed:
Coleman, P.K. (2006). Resolution of unwanted pregnancy during adolescence through abortion versus childbirth: Individual and family predictors and consequences. <i>Journal of Youth and Adolescence</i> , 35, 903-911.	One hundred and thirty adolescents in grades 7-11 who completed both Waves I & II and experienced a Pg described as "not wanted" or "probably not wanted".	AB N= 65 DEL N=65	Single-item measures of counseling, 12 month trouble sleeping, 30 day cigarette use, 30 day marijuana use, 12 month alcohol use, problems with parents and with school due to alcohol use.	Controlling for risk taking and desire to leave home, AB group more likely to have counseling, trouble sleeping, and use marijuana in past 30 days (problems with parents due to alcohol use approached significance).	Number of total Pgs unknown, but small ns raise questions about underreporting and drop-out rates. 1-item outcome measures psychometrically weak. Percentages and ns for outcome variables not reported so frequency of problem unknown; previous mental health problems not controlled. Given the large number of variables in the data set, why these particular variables were included is unclear. Not clear when counseling occurred.

ates in analyses of that outcome variable. Analyses were often based on small subgroups or subgroups for which no sample size was provided. On the other hand, the overall large sample sizes used for some analyses mean that small effects that are statistically significant may have little clinical significance.

Although initially based on a national probability sample, the ability to assess prevalence of mental health problems among women who have abortions from this data set is limited because (1) abortion has been under-reported in the NLSY compared with national norms;

(2) sample weights, required to construct population estimates from the data, were not used in the analyses of any of the studies; and (3) the measurement of mental health outcomes was limited to self-esteem, depression risk, and substance abuse. No actual measures of psychopathology were included.

The potentially strongest designs focused on mental health outcomes associated with unintended first pregnancy. However, the practices of excluding women who became pregnant at a young age (before 1979 or 1980) and differentially excluding women having

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Results	Notes and Additional Limita- tions Specific to Study Listed:
Hope, T.L., Wilder, E.I., & Watt, T.T. (2003). The relationships among adolescent pregnancy, pregnancy resolution, and juvenile delinquency. <i>Sociological Quarterly</i> , 44, 555-576.	ADD-HEALTH data from Waves I & II used to examine the relationships among adolescent Pg, Pg resolution, and delinquent behavior. Women who experienced Pg prior to Wave I, miscarried, or were still Pg at Wave II excluded; 360 ever Pg adolescents who had an abortion or kept baby and did not choose adoption were identified. Longitudinal analysis based on 156 women who became Pg between Waves I & II reported here. Although adoption grp had sig higher delinquency rate than Kept baby group, the small n (4), precluded inclusion in longitudinal analyses.	Longitudinal analysis: AB N=87 Kept baby N=69	Comparing AB vs. Kept baby groups, 3.6% vs. 15.0% on welfare; 39.9% vs. 23.9% in intact families. These variables not controlled. Most relevant here: longitudinal analyses of relationship between Pg outcome & cigarette smoking or marijuana use on at least 1 day in the past 30 days.	AB grp reported higher rates of cigarette smoking and marijuana use than those who kept baby both prior to their Pg (Wave I) and subsequent to their Pg (Wave II). Keeping baby associated with a decrease in cigarette and marijuana use after Pg; no sig change in such use was found before vs. after AB grp.	Number of total Pgs unknown, but small ns raise questions about underreporting and drop-out rates that may advantage Kept baby group; measures psychometrically weak and of unknown clinical significance. Percentages and ns for outcome variables not reported, so bases for % of problems in various grps unclear. The extent to which delinquent mothers may have higher drop out rates than other mothers is unknown. Although adoption grp not analyzed due to low n, the sig higher overall rate of delinquency for that grp emphasizes importance of recognizing heterogeneity in women who deliver.

abortions subsequent to first pregnancy from the delivery group but not the abortion group were shown to bias results toward overestimating adverse effects of abortion in this data set. In the one study focusing on first pregnancy that did not use differential exclusion and was based on codes provided by NLSY staff, the proportion of women who met or exceeded the CESD cutoff scores did not significantly differ between abortion (25%) and delivery (28.3%) groups (Schmiege & Russo, 2005).

Washington, DC, Metropolitan Area Drug Study.
 Coleman, Reardon, and Cogle (2005) used this public release data set to examine substance use during pregnancy as a function of reported reproductive history.

The initial sample, which consisted of 1,020 women interviewed after giving birth in Washington, DC, area hospitals in 1992, was predominantly never married, Black, of low socioeconomic status, and oversampled for low birth weight and preterm infants, and self-reported drug use. Of these cases, Coleman et al. (2005) selected those who in their interview reported no abortions, one abortion, or multiple abortions prior to their recent pregnancy and examined their reported drug use during their recent pregnancy (see Table 2). Adjusted for age, income, and number of people living in the house, a statistically higher odds ratio was reported for the use of legal and illegal substances during the index pregnancy if the woman had reported one prior abortion compared with no abortions, but not if she had

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

National Survey of Family Growth (NSFG)

General Description: The NSFG Cycle V sample is a subsample of 10,847 women aged 15-44 drawn from the larger national probability sample of the National Health Interview Survey. The NSFG is thus based on a complex stratified, multistage design that requires using sampling weights in computing statistics.

Limitations Common to All Studies Based on this Data Set: Retrospective self-report data that may involve recall of precise timing of key variables (e.g., abortion, onset of anxiety symptoms), occurring decades previously.

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Additional Limitations Specific to Study Listed:
Cougle, J., Reardon, D.C., Coleman, P.K., & Rue, V. M. (2005). Generalized anxiety associated with unintended pregnancy: A cohort study of the 1995 National Survey of Family Growth. <i>Journal of Anxiety Disorders</i> , 19, 137-142.	Study sample: 1. All women having an unintended Pg ending in abortion for their first Pg event. 2. All women having an unintended Pg ending in live birth delivery for their first Pg event who had no abortions after that Pg. Women who experienced a prolonged period of anxiety previous to or at the same age as the Pg event were excluded from the sample.	AB N=1033 DEL N=813	Dichotomous measure (yes/no) of anxiety symptoms.	Sig higher rate of anxiety symptoms in AB vs. DEL group (13.7% vs. 10.1%), controlling for race and age at interview. In stratified sub-analyses, difference sig for unmarried or women under 20 at 1st Pg, but not for married women.	Women reporting pre-Pg anxiety excluded so cannot generalize to all first unintended pregnancies; misleading language implies generalized anxiety disorder (GAD) is assessed, but items used to construct anxiety variable are not congruent with DSM definitions of generalized anxiety disorder, making clinical implications problematic; differential exclusion of women with subsequent abortions from DEL but not AB grp; sampling weights not used in statistical analyses; stratification used rather than controlling for relevant variables. No attempt to control for any violence history although questions re rape experience available in data set.

reported multiple abortions compared with no abortions (with the exception of use of cigarettes during pregnancy). Notably, these analyses did not control for history of drug use prior to the pregnancy. They also did not control for the wantedness of the pregnancy, although those data were available in the data set. Because this study is based on a specialized sample, estimates of mental health problems among women in the

United States who have an abortion cannot be determined from this study.

National Pregnancy and Health Survey. Coleman, Reardon, Rue, and Cougle (2002a) used data from this survey conducted in 1992 to examine the association between retrospective reports of a previous abortion and use of alcohol, cigarettes, or illicit drugs

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Commonwealth Fund Health of American Women Survey

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Results	Additional Limita- tions Specific to Study Listed:
Russo N.F., & Denious J. (2001). Violence in the lives of women having abortions: Implication for practice and public policy. <i>Professional Psychology: Research and Practice</i> , 32, 142-150.	Secondary analyses of a random household telephone survey of over 2,500 women and 1,000 men aged 18 or over and residing in the continental U.S., conducted in 1993. Analyses based on responses of 2,525 women, 324 of them identified as having had at least 1 abortion; ns varied depending on missing data.	AB N=324 Others N=2,201	Global self-esteem (RSE); abbreviated CES-D (6 items); 1-item measures of suicidal ideation in past year; if told by doctor she had anxiety/depression in past 5 years, 1-item life satisfaction measure.	AB correlated positively with CES-D (.08), having suicidal thoughts (.08), being told by a doctor had anxiety/depression (.08) & negatively with life satisfaction (-.06). Also correlated with experiencing rape (.06), childhood physical (.15) & sexual (.18) abuse, having a violent partner (.11) & a partner who refused to use condom (.06). Controlling for race, education, children living at home, marital status, and partner and violence variables, abortion not sig related to any outcome variable.	Outcome and violence measures psychometrically weak. Timing of events vis-à-vis abortion unknown. Abbreviated CES-D used; Only women married or living as a couple were asked about partner violence. Limited generalizability of study group: have telephone, younger teenagers not included; older age (median 40-44); 57% married. Low reported abortion rate (13%) could reflect underreporting and/or recall bias. Only one question asked about abortion history; repeat abortions not identified. Comparison is with other women, not women with unintended Pg.

during the most recent pregnancy. The initial sample consisted of 2,613 women who participated shortly after giving birth in hospitals within the United States. The women wrote down answers in response to interviewer questions; responses were concealed from the interviewer. Samples selected for analysis were limited to three groups who had recently given birth: women with one previous pregnancy resulting in an induced abortion ($n = 74$), women with one previous pregnancy resulting in live birth ($n = 531$), and women with no previous pregnancies ($n = 738$). The majority of the women were White, married, and employed full-time. Dichotomous measures of drug and alcohol use during most recent pregnancy were used as outcome variables. Analyses revealed that women

who reported a previous abortion also reported higher rates of any illicit drug use, marijuana use, and alcohol use than did women who had one previous live birth or were first-time mothers. The researchers adjusted for sociodemographic covariates by stratifying those related to substance use outcomes and conducting separate analyses for each level of these variables. Although these analyses identified some differences in the relationship of reproductive history to alcohol and drug use for different levels of marital status, income, and other demographic variables, findings are suspect because of the small number of participants in the abortion group and the failure to correct for the relatively large number of significance tests. Other limitations include the absence of controls

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Harvard Study of Moods and Cycles

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Additional Limitations Specific to Study Listed:
Harlow, B.L., Cohen, L., Otto, M.W., Spiegelman, D., & Cramer, D.W. (2004). Early life menstrual characteristics and pregnancy experiences among women with and without major depression: The Harvard study of moods and cycles. <i>Journal of Affective Disorders</i> , 79, 167-176.	Subsample drawn from a cross-sectional sample of 4,161 women between 36-45 years of age residing in seven Boston metropolitan area communities consisting of 332 women who had a past or current history of major depression as measured by DSM criteria and 644 women with no such history.	Comparisons made between 332 depressed and 644 nondepressed women.	Percentage of women who reported experiencing at least one abortion for depressed (DEP) and nondepressed (NDEP) groups.	Percentage of women having had at least one abortion 34.1% and 24.1%, for DEP & NDEP grps, respectively; higher % of abortions in the DEP group reflected a higher % of women having multiple abortions (14.8% vs. 6.2%). Controlling for age, age at menarche, educational attainment, and marital experience, no sig differences between % of women with a lifetime history of dep (19.3%) and no history of dep (17.9%) reporting at least one abortion. Women with lifetime history of major dep upon study enrollment were 3 times more likely to report having had multiple abortions before their first onset of depression than were nondepressed women. Also found a strong association between dep and marital disruption.	Direct comparisons between women reporting abortion vs. delivery were not conducted. Wantedness of Pg not assessed. Association between dep and marital disruption underscores importance of controlling for marital status when seeking to assess the independent contribution of abortion to depression risk. Retrospective reproductive history and depression onset data. Researchers suggest variety of unassessed antecedent conditions may underlie results, including involvement in abusive relationships.

for wantedness of the recent pregnancy, history of drug use prior to the pregnancy, or previous mental health.

Fertility and Contraception Among Low-Income Child Abusing and Neglecting Mothers in Baltimore, MD, 1984-1985 (Baltimore Study). Coleman, Maxey, Rue, and Coyle (2005) analyzed this data set to examine the association between self-reported abortion or miscarriage/stillbirth history and child

abuse and/or neglect, as identified by Child Protective Services. The purpose of the original study had been to study family patterns and contraceptive use among maltreating mothers. Samples of 118 physically abusive mothers, 119 neglecting mothers, and 281 mothers without maltreatment offences were selected from a sample of 518 mothers who were receiving Aid to Families With Dependent Children (79.9% Black and 93.2% unemployed). In an in-home interview, 159 of these women reported having

Table 2: Secondary Analyses of Survey Data—Abortion vs. Comparison Groups
U.S. STUDIES (continued)

National Pregnancy and Health Survey

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Results	Additional Limita- tions Specific to Study Listed:
Coleman, P.K., Reardon, D.C., & Cougle, J. (2002). A history of induced abortion in relation to substance use during subsequent pregnancies carried to term. <i>American Journal of Obstetrics and Gynecology</i> , 187, 1673-1678.	Data drawn from the National Pregnancy and Health Survey conducted in 1992 whose purpose was to assess drug and alcohol consumption in a national sample of pregnant women (N= 2,613). Hospitals with < 200 annual births were selected in the first stage of sample selection; individual mothers within hospitals were randomly selected in the second stage. Soon after delivery, women were interviewed about reproductive history and completed a drug-use questionnaire answer sheet in response to interviewer questions. Samples used in analyses were limited to women who recently had given birth and had one previous induced abortion, one previous birth, or no previous births or abortions.	The primary sample of women with a recent delivery (n = 607) had two subgroups: 74 women with one previous induced abortion and 531 women with one previous birth. The secondary sample included 738 first-time mothers with no previous abortions. Both grps were primarily White, married and employed full-time. Average age of the two grps was 26.5 and 23.4 yrs, respectively.	Association between previous reproductive outcome and usage of alcohol or illicit drugs during most recent pregnancy. Differential odds ratios for use of any illicit drugs, marijuana, cigarettes and alcohol reported for 1 previous abortion vs. 1 previous birth group, and 1 previous abortion vs. first birth group. Adjusted for covariates by stratifying covariates related to substance use type and running separate analyses.	Women with a previous abortion had higher rates of any illicit drug use, marijuana use and alcohol use, than women with a previous live birth. Differences between reproductive history groups appeared greater when time since previous pregnancy was longer (3-5 vs. < 2 years). The abortion group also reported higher rates of illicit drug use, marijuana, and alcohol use than first-time mothers.	Samples analyzed not representative of total NPHS sample or of U.S. women giving birth. Retrospective self-reports of abortion may be unreliable. Abortion likely underreported. Single-item outcome measures. No statistical adjustment for number of significance tests. Confounds not controlled. Small size of abortion group led to many cell counts <5 in subgroup analyses which were intended to control for confounds. Rates of use not reported. Differences found could be due to other unmeasured factors such as whether pregnancy intended, partner violence, or sexual abuse. Comparisons between previous abortion and previous birth groups could be explained by child care demands on mothers or differential stress of first vs. later completed pregnancy.

had at least one abortion, and 133 reported at least one miscarriage or stillbirth (both occurring on average 6-7 years earlier). Controlling for a large number of single-item covariates found in preliminary analyses to be associated with maltreatment (and that varied depending on their association with the outcome variable, e.g., education was controlled only in the analyses on physical abuse; employment controlled

only in the analyses on neglect), women reporting one abortion were not more likely than those reporting no abortions to be in the child neglect group but were significantly more likely to be in the physical abuse group. History of multiple induced abortions, however, was not related to increased risk for either abuse or neglect. In contrast, maternal history of multiple miscarriages and/or stillbirths compared

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Washington, DC, Metropolitan Area Drug Study

Citation	Sample & Procedure	Sample Sizes	Primary Outcome	Key Findings	Additional Limitations Specific to Study Listed:
Coleman, P., Reardon, D. C., & Cougle, J. R. (2005). Substance use among pregnant women in the context of previous reproductive loss and desire for current pregnancy. <i>British Journal of Health Psychology</i> , 10, 255-268.	Data drawn from the public release data set of the Washington DC Metropolitan Area Drug Study (CD*MADS). The initial sample, constructed to oversample for low birth weight, pre-term, and admitted maternal drug use, consisted of 1020 women giving birth in Washington, DC area hospitals in 1992. The initial sample was predominantly never married, Black, between 19 and 34 years of age, high school or less education, and of relatively low family income (under \$20,000). Of these cases, those with known medical outcomes of previous pregnancies were selected for further analysis.	Sample sizes varied across analyses. Key comparisons in Table 3, in which odds ratios for drug use during the current Pg as a function of abortion history, appear based on 144 women reporting no prior abortions vs. 282 women reporting one or more abortions prior to the index delivery. [These # not directly reported in paper but were determined through examination of the public release data set used in these analyses. Numbers essentially consistent with %d & methods reported in the paper.]	Differential odds ratios for the use of marijuana, cigarettes, alcohol, crack cocaine, other cocaine, and any illicit drugs are reported for 1 previous abortion vs. no abortion history and 2 or more abortions vs. no abortion history after statistical adjustment for number of prior births, miscarriages, and still births; age; education; number of people the respondent lives with; and a binary indicator reflecting if prenatal care was sought in the first trimester.	Adjusted for covariates, a statistically higher odds ratio was reported for the use of legal and illegal substances during the index pregnancy if the woman had a prior history of abortion.	The sample very specialized. No indication that sampling fractions used in analysis to reweight sample. Rates of use not reported for comparison grps. Many of the illegal substance categories are fairly rare (e.g. there are only 58 cases of any reported crack cocaine use during Pg among the subset of cases who had usable data on abortion history). Results look very different for covariate-adjusted analyses and unadjusted analyses. Intendedness of Pg not used as co-variate in abortion analyses.

with no history was associated with increased risk of both child physical abuse and neglect. Because this study is based on a highly specialized sample, findings cannot be generalized to the population of women in the United States.

Health of American Women Survey. Russo and Denious (2001) used data from this survey, sponsored by the Commonwealth Fund, to examine correlations among abortion history, violence history, and mental health outcomes. This telephone survey was based on a national sample of men and women 18 years of age or older, with oversampling of ethnic minorities. Among the 2,525 women surveyed, 324 reported

having had an abortion to the interviewer. Compared with other women, a larger percentage of women in the abortion group reported experiencing suicidal thoughts in the past year and having a doctor give them a diagnosis of anxiety or depression in the past 5 years. Having an abortion was also slightly but significantly correlated with higher depressive symptoms and lower life satisfaction. When violence history and relevant demographic and partner variables were controlled, however, abortion was no longer significantly related to diagnoses of depression or anxiety, CES-D score, or the life satisfaction measure. This study, like the others of this type, has several limitations. Abortion history was assessed through self-report (in this

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
U.S. STUDIES (continued)

Fertility and Contraception Among Low Income Child Abusing and Neglecting Mothers in Baltimore, MD, 1984-1985

Citation	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Results	Additional Limita- tions Specific to Study Listed:
Coleman, P., Maxey, C.D., Rue, V.M., & Coyle, C.T. (2005). Associations between voluntary and involuntary forms of perinatal loss and child maltreatment among low income mothers. <i>Acta Paediatrica</i> , 94, 1476-1483.	Data drawn from Fertility and Contraception among Low Income Child Abusing and Neglecting Mothers in Baltimore, MD, 1984-1985, a study of family patterns and contraceptive use among maltreating mothers. Sample of 518 mothers (age range 18-50; 79% Black; 6.8% employed) who were receiving AFDC. All women interviewed in home.	One hundred and eighteen physically abusive mothers and 119 neglecting mothers selected from cohort receiving child protective services (CPS) and 281 mothers without maltreatment offences. In interview, 100 women reported 1 abortion, 59 reported 2+ (abortion ave 6.5 years earlier), 99 reported 1 miscarriage or stillbirth, 34 reported 2+ (ave 7.1 yrs earlier).	Association between self-reported abortion or miscarriage/stillbirth history and being in the physically abusing or neglecting groups. Logistic analyses controlled for covariates (single-item measures) associated with maltreatment (e.g., more children, history of depression, worries about income, etc).	Adjusted for covariates, women reporting 1 abortion were not more likely than those reporting no abortions to be in child neglect group, but were sig more likely to be in physical abuse group. History of multiple induced abortions not related to increased risk for either abuse or neglect. Maternal history of multiple miscarriages and/or stillbirths compared to no history was associated with increased risk of physical abuse and neglect.	Retrospective self-reports of abortion in interview unreliable. Abortion likely underreported. Sample not representative of U.S. women. No info about nature of abortion. Single-item measures of covariates. Causal direction ambiguous. Same factors (e.g., poverty, drug use) may contribute to increased risk of child maltreatment and abortion. Intendedness of Pg not assessed, and given the poor health among this study population, lack of information about whether the previous abortion was for therapeutic reasons is a particular limitation.

case over the phone), and the rate of reported abortions was low compared with national norms, raising concerns about biases associated with underreporting. It cannot be determined from this data set whether the abortion took place before or after the violence occurred, or whether diagnoses of anxiety or depression occurred pre- or post abortion. In addition, sampling weights were not used.

National Survey of Family Growth (NSFG). Cougle et al. (2005) used data from the 1995 NSFG to examine the association between outcome of first- unintended pregnancy (abortion vs. delivery) and an occurrence of "generalized anxiety" lasting more than 6 months defined by a cutoff score). All variables—reproductive history, episodes of anxiety, as well as the timing of those episodes with respect to pregnancy—were deter-

mined retrospectively via self-reports, raising questions about reliability and underreporting of abortion. As in their earlier studies, women with subsequent abortions were differentially excluded from the delivery group but not the abortion group. Controlling for race and age at interview, women in the abortion group were more likely to be classified as having had an episode of generalized anxiety postpregnancy than women in the delivery group (13.7% vs. 10.1%). Sample weights were not used, so these percentages cannot be used for normative estimates. Although information on rape history, known to be related to both unintended pregnancy and anxiety, was in the data set, it was not controlled. The anxiety items were not congruent with the DSM definition of generalized anxiety disorder, raising questions about the clinical significance of the outcome variable.

Table 2: Secondary Analyses of Survey Data – Abortion vs. Comparison Groups
INTERNATIONAL STUDIES–NEW ZEALAND

Christchurch Health and Development Study

General Description: The NSFG Cycle V sample is a subsample of 10,847 women aged 15-44 drawn from the larger national probability sample of the National Health Interview Survey. The NSFG is thus based on a complex stratified, multistage design that requires using sampling weights in computing statistics.

Limitations Common to All Studies Based on this Data Set: Retrospective self-report data that may involve recall of precise timing of key variables (e.g., abortion, onset of anxiety symptoms), occurring decades previously.

New Zealand	Data Source/ Population Studied	Sample Sizes	Primary Outcome	Results	Additional Limitations Specific to Study Listed:
Fergusson D.M., Horwood, L.J., & Ridder, E. M. (2006). Abortion in young women and subsequent mental health. <i>Journal of Child Psychology & Psychiatry</i> , 47, 16-24.	The Christchurch Health and Development Study is a longitudinal study of a cohort of 1,265 children born in 1977 in the Christchurch, New Zealand, urban region who were studied from birth to age 25, including 630 females; 41% of women Pg on at least one occasion prior to age 25; 14.6% had at least one abortion. Sample sizes in analyses ranged from 506 and 520 depending on the timing of assessment. Details on Ns for prospective analyses were provided in personal communication from the author.	Concurrent analyses: AB N= 74 DEL N= 131 Never Pg N= 301 Prospective analysis: AB N=48 DEL N= 77 Never Pg N= 367	DSM-IV mental disorders (including major dep, overanxious disorder, GAD, social phobia, & simple phobia, and suicidal behavior for intervals 15-18, 18-21 and 21-25 years, controlling for childhood, family, and related confounding factors. Outcomes for concurrent analyses: yes/no diagnosis of major dep, anxiety disorder, alcohol and illicit drug dependence, suicidal ideation in previous 12 mo, and total # of disorders; in prospective analysis, total number of disorders from 21-25 yrs.	In concurrent analyses, controlling for covariates, AB grp had sig (p<0.05) higher rates of depression, suicidal ideation, illicit drug dependence, & total mental health problems than the DEL grp & except for alcohol and anxiety disorder, significantly higher rates of disorder than the Never PG grp. A prospective analysis used Pg/abortion history prior to age 21 to predict mental health outcomes from 21-25 years. Similarly, after covariate adjustment, the AB grp had a sig higher total # of disorders than the other grps, which did not sig differ from each other.	Neither intendedness nor wantedness of Pg controlled; screening criteria related to mental health for legal abortion in New Zealand may bias portrait of outcomes. Abortion is underreported. N too small for multiple abortions to be analyzed separately. Although a longitudinal study, most results reported involved the concurrent assessment of Pg status and mental health. The prospective analysis was limited to number of disorders owing to the relatively sparse data for specific disorders over the interval 21-25 years and the small number of women who became pregnant by age 21.

Notes: AB = Abortion DEL = Delivery, Pg = pregnancy; ACOG = American College of Obstetricians and Gynecologists; ICD = International Classification of Diseases; Grp = Group; Sig = Significance

National Longitudinal Study of Adolescent Health (ADD-Health). Two studies were based on the ADD-Health data set, a longitudinal, nationally representative, school-based survey of adolescents. Coleman (2006a) analyzed data from the ADD-Health to examine the relationship between reproductive history and various problems in adolescents. From a much larger sample of students who had completed an in-school questionnaire at Wave I (N = 90,118) and a computer-

assisted home interview at Wave II (N = 12,105), Coleman selected adolescents in grades 7 through 11 who had completed both Wave I and Wave II and who reported experiencing a pregnancy they described as “not wanted” or “probably not wanted” that was resolved through abortion (n = 65) or delivery (n = 65). She then examined the likelihood that adolescents who reported abortion versus delivery also reported receiving counseling for psychological or emotional

problems, having trouble sleeping during the past year, using cigarettes or marijuana during the past 30 days, using alcohol during the past year, or reported having problems with parents because of alcohol use. All outcomes were assessed with single-item measures. Adjusted for covariates previously shown to differ between the two groups (risk-taking and desire to leave home), girls who reported an abortion were more likely than girls who delivered to say they had ever had counseling, trouble sleeping during the past year, and used marijuana in past 30 days. No differences were observed on frequency of alcohol use or cigarette smoking.

Strengths of this study included the use of a comparison group of girls who delivered unwanted pregnancies, the weighting of design factors in the analyses, and efforts to enhance the accuracy of self-reports of sensitive topics (respondents listened to prerecorded questions through earphones and entered their own answers). Nonetheless, problems of sampling and measurement limit the utility of this study. The extremely small number of girls in the eventual sample analyzed ($N=130$), especially given the very large original sample (of approximately 6,000 girls), raises questions about underreporting, drop-out rates, and exclusion criteria. Given that the sample is school-based, adolescents who drop out of school to care for a child would not be included in the study. The single-item measures of psychological problems are psychometrically weak and clinically suspect. Because the percentages and N s for outcome variables were not reported, the frequency with which problems occurred cannot be determined. Furthermore, the measure of counseling asked whether the respondent had ever received counseling for psychological or emotional problems—it cannot be determined from this item whether counseling occurred prior or subsequent to the pregnancy.

Hope, Wilder, and Watt (2003) used data from the ADD-Health study (Waves I and II) to examine the relationships among adolescent pregnancy, pregnancy resolution, and delinquent behavior. Although delinquency includes behaviors that are not part of the mental health focus of this review (e.g., lying to parents/guardian, taking part in a fight), one domain of delinquent behavior examined (alcohol use, use of illegal substances) is within the purview of this review. Thus, we focus here on longitudinal analyses examining the relationship between pregnancy resolution and

substance use. In a set of prospective analyses focusing on adolescent girls who became pregnant between Wave I and II of the survey, Hope et al. examined the relationship of pregnancy resolution (abortion vs. kept baby) to reports of having smoked cigarettes or marijuana at least 1 day in the past 30 days. These comparisons of the abortion and “kept baby” groups excluded girls who experienced pregnancies prior to Wave I as well as those who miscarried or were still pregnant at Wave II.

Young women who had abortions reported higher rates of cigarette smoking and marijuana use than young women who kept their baby, both prior to their pregnancy (Wave I) and subsequent to their pregnancy (Wave II). Keeping the baby was associated with a decrease in reported cigarette or marijuana use between the two waves of data collection, leading the authors to conclude that adolescent motherhood functions as a social control on delinquent behavior. In contrast, having an abortion was not associated with a change in rates of smoking or marijuana use from Wave I to Wave II, leading the authors to conclude that terminating a pregnancy through abortion does not increase the likelihood of delinquent behavior or substance use.

In addition to strengths and weaknesses of the ADD-Health school-based database described above, this study is limited by single-item measures of cigarette and marijuana use that are psychometrically weak. Furthermore, despite the large initial sample size of over 6,000 girls, the number of pregnant girls (69 who had abortions, 87 who kept their baby) in the final sample was small.

The Harvard Study of Moods and Cycles. Harlow, Cohen, Otto, Spiegelman, and Cramer (2004) used data from a cross-sectional sample of 4,161 women between 36-45 years of age residing in the Boston metropolitan area to examine the relationship of early life menstrual-cycle characteristics and reproductive history to onset of major depression later in life. They analyzed data from a subsample of 332 women who met DSM criteria for having had major depression and 644 women with no current or past history of major depression. In-person interviews were used to establish mental health status and to gather information on demographic and lifestyle characteristics, menstrual and reproductive history, past and current